REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-02-
0134

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
March 16, 2002	Final	971231-981231

4. TITLE AND SUBTITLE
Wavelet modulation in fading channels

Optimal Detection and Signalling in Fast Fading Channels

5a. CONTRACT NUMBER
F49620-98-1-0093

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Chit Lo and Todd K. Moon

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Utah State University
Electrical and Computer Engr.
4120 Old Main Hill
Logan, UT 84322

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR
801 N. Randolph Street
Room 732
Arlington, VA 22203-1977

20020419 167

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release,
distribution unlimited

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
NOTICE OF TRANSMITTAL DTIC. THIS TECHNICAL REPORT
HAS BEEN REVIEWED AND IS APPROVED FOR PUBLIC RELEASE
LAW AFR 190-12. DISTRIBUTION IS UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We characterize optimal detection and signaling in terms of signal-to-noise for a fast or frequency-selective channel when the autocorrelation function of the channel is known. Our result characterizes the optimization problem as an eigenvalue/eigenfunction problem with integral operators. For detection, the well-known matched filter and RAKE receivers are limiting cases of our results. In addition, we provide a closed-form expression for the probability density and the cumulative probability function of a quadratic form for zero mean complex gaussian vectors, which is used to evaulate the probability of error of quadratic receivers. Additionaly, we provide a procedure that finds pairs of signals or filters for digital communication over a fading channel. With the land-mobile fading channel, the signal/filter pairs obtained with our procedure have a performance significantly better than that of the traditional flat-top pulse and a raised cosine pulse in a PPM pair in terms of the probability of error rate.

15. SUBJECT TERMS
Wavelet modulation; fading channels; eigenfunction signal design.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Todd K. Moon
unclassified	unclassified	unclassified		xi+121	19b. TELEPHONE NUMBER *(include area code)* 435 797 2970

Optimal Detection and Signaling in Fast Fading Channels

Chit Lo

Todd K. Moon

Electrical and Computer Engineering Department
Utah State University
Logan, UT 84322-4120

Abstract

We characterize optimal detection and signaling in terms of signal-to-noise for a fast or frequency-selective channel when the autocorrelation function of the channel is known. Our result characterizes the optimization problem as an eigenvalue/eigenfunction problem with integral operators. For detection, the well-known matched filter and RAKE receivers are limiting cases of our results. In addition, we provide a closed-form expression for the probability density and the cumulative probability function of a quadratic form for zero mean complex Gaussian vectors, which is used to evaluate the probability of error of quadratic receivers. Additionally, we provide a procedure that finds pairs of signals or filters for digital communication over a fading channel. With the land-mobile fading channel, the signal/filter pairs obtained with our procedure have a performance significantly better than that of the traditional flat-top pulse and a raised cosine pulse in a PPM pair in terms of the probability of error rate.

Contents

List of Tables ix

List of Figures x

1 Introduction 1
- 1.1 General Background 1
- 1.2 Literature Review 3
- 1.3 Methodology 7
- 1.4 Contributions to the Field of Knowledge 8
- 1.5 Dissertation Structure 9

2 Channel Models 11
- 2.1 Introduction 11
- 2.2 Fading Channel Model 11
- 2.3 Categories of Degradation 14
- 2.4 Model for Frequency-Nonselective Fading Channels 15
- 2.5 Model Used in This Dissertation 16

3 Optimal Detection and Signaling 18
- 3.1 Optimal Detection 18
- 3.2 Characterization of Optimal Detectors 19
- 3.3 Optimal Signaling 22
- 3.4 Characterization of Optimal Signaling for Single-Path Fading 22

4 Iterative Algorithms for Finding Solutions 28
- 4.1 Iterative Algorithms 28
- 4.2 Convergence 31
- 4.3 Introduction of Frequency Constraints 33
- 4.4 Signal Sets 35

5 The Study of Probability of Error 39
- 5.1 Introduction 39
- 5.2 Quadratic Forms of Complex Normal Vectors 39
- 5.3 Cumulative Distribution Function 42
- 5.4 A Partition of Unity 43
- 5.5 Quadratic Receiver 44

6 Implementation of the Algorithms **50**
6.1 Conversion of a Linear Integral Equation to Its Discrete Equivalent 50
6.2 Discrete Version of Algorithm 2 and 3 52
6.3 Multipath Fading Channel Model 55
6.4 Numerical Methods 56

7 Results and Observations **61**
7.1 Different Initial Vectors 61
7.2 Frequency Containment 61
7.3 Effect of Multipath Channels 72
7.4 Effect of Synchronization Uncertainty 72
7.4.1 Fast Fading 72
7.4.2 No Fading 72
7.5 Signal Sets 77
7.5.1 CCPs 77
7.5.2 PPM with CPs 84
7.6 The Convergence of Algorithms 87

8 Conclusion and Discussion **88**
8.1 Summary 88
8.2 Future Work 90

References **92**

Appendixes **101**

A Derivative of a Functional with Respect to a Function **102**

B Inverse of a Symmetric Kernel **105**

C Programs **106**

List of Tables

Table Page

2.1 Spectra and Autocorrelation Function of Fading Process. 14

2.2 Categories of Degradation According to Different Fading Phenomena. . . . 15

6.1 Delay and Amplitude Profile of GSM Channels. 55

List of Figures

Figure Page

2.1 System model of a fading channel. 16

4.1 Procedure for generating signal/filter pairs. 36

5.1 Contour map. 41

5.2 Detection unit of quadratic receiver. 45

6.1 The illustration of windowing effect. 55

6.2 Approximation of $J_0(t)$ with a positive sequence. 56

6.3 The illustration of the result from Algorithm 4 with 8192 data points. 59

7.1 Illustration of the effect of different initial vectors I. 62

7.2 Illustration of the effect of different initial vectors II. 63

7.3 Single path, perfect synchronization, $\beta = 0.1$ and 0.3. 64

7.4 Single path, perfect synchronization, $\beta = 0.5$ and 0.7. 65

7.5 Single path, Gaussian distribution with $\sigma = 1.124$, $\beta = 0.1$ and 0.3. 66

7.6 Single path, Gaussian distribution with $\sigma = 1.124$, $\beta = 0.5$ and 0.7. 67

7.7 HT, perfect synchronization, $\beta = 0.1$ and 0.3. 68

7.8 HT, perfect synchronization, $\beta = 0.5$ and 0.7. 69

7.9 TU, uniformly distributed, $\mathcal{U}(-4.15\times10^{-6}, 4.15\times10^{-6})$sec, $\beta = 0.1$ and 0.3. 70

7.10 TU, uniformly distributed, $\mathcal{U}(-4.15\times10^{-6}, 4.15\times10^{-6})$sec, $\beta = 0.5$ and 0.7. 71

7.11 Effect of multiple paths I. 73

7.12 Effect of multiple paths II. 74

7.13 Effect of synchronization uncertainty I. 75

7.14 Effect of synchronization uncertainty II. 76

7.15 CP for channels without fading I. 78

7.16 CP for channels without fading II. 79

7.17 CP for channels without fading III. 80

7.18 CCPs for single path, perfect synchronization $\beta = 0.1$ and $\beta = 0.3$. 81

7.19 CCPs for single path, perfect synchronization $\beta = 0.5$ and $\beta = 0.7$. 82

7.20 SNR vs. $P(e)$, CCPs for single path, perfect synchronization. 83

7.21 TU, uniformly distributed, $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec. 85

7.22 HT, perfect synchronization. 85

7.23 Single path, Gaussian-distributed with $\sigma = 1.124$. 86

7.24 Single path, perfect synchronization. 86

Chapter 1
Introduction

1.1 General Background

The recent explosion in internet traffic proclaims the arrival of the Information Age [1]. This exploding internet traffic volume at the same time imposes a tremendous challenge on communication engineers to meet the unlimited demand for information bandwidth, both cable based and wireless [2]. On the other hand, this demand also brings us opportunities that we were not able to imagine before. A recent prediction asserts that "data" communication will soon overtake "voice" communication [3], and future development is heading toward personalization and mobilization [4]. A wireless "information superhighway" is envisioned to fulfill the demand for multimedia networks with multi-service requirements [5–7].

At present, systems of wireless communication, mainly cellular, are only able to provide low-speed data communication at error rates that are far from acceptable in wireless network connections. This is because these systems have been established based on technology distant from theoretical limits [8]. One of the main reasons for this drawback in technology is the hurdle of multipath fading channels [9, 10]. A very large proportion of the data bits is used for signal integrity, either by extensive code correction or multiple repetition. The cellular system we use today is designed more to circumvent the problems imposed by the multipath fading phenomenon than actually to solve them. Even if there are efforts to overcome the problems of multipath-fading channels, those techniques applied are relatively primitive. A typical example is to repeatedly transmit a message through a channel so that a correct reception of the message can be determined by a majority vote (e.g., the United Kingdom total access communications system repeats its message 11 times [11]). Also, to avoid excessively rapid fluctuation of the received signal power

caused by fading, a specific signaling rate is chosen according to the rate of fading so that the received signal can be regarded as approximately constant over a data symbol interval (i.e., the slow fading assumption). At the same time, the signaling rate is limited by severe problems with inter-symbol interference (ISI), which is a frequency selective problem due to multipath transmission [12]. However, there is an indication that high-speed reliable communication is possible if the user is willing to optimize what the media could provide [13]. Because of this belief, considerable research effort has been devoted in the area of multipath-fading channel communication for the past five decades [9, 14–22].

Usually multipath-fading channels are merely called fading channels [23]. However, there are actually two basic degradations, i.e., multipath and deep fades. Multipath, which results from reflections of the transmitted signal from reflective surfaces between the transmitter and the receiver, causes delayed and scaled versions of signals to be superimposed at the receiver. When the arrival times of the different rays are of the same order of magnitude as the duration of the transmitted signals, successive signals are smeared together, thus resulting in ISI. The span of the excess delay (i.e., the time between the first and last received components during which the received signal power falls to some threshold level below that of the strongest component) is directly related to the physical distortion of the signal by the channel. This kind of distortion varies according to the difference in frequency contain. Frequency selectiveness is the term we use when the distortion rate, the channel gain, and the phase variation change significantly for a small variation in frequency. When the arrival time difference is comparable to the period of the carrier frequency, deep fades result. Deep fades are the phenomena that waves of different phases superimpose and interfere constructively or destructively. This kind of interference may cause an extremely low received signal power. Sometimes, this deviation of signal power can be more than 40 dB [9]! The distance between nulls is approximately 0.33 meter for 900 MHz cellular systems. Deep fades can also result from a relative motion between the transmitter and the receiver, which also results in superposition interference called the Doppler effect [24]. At very high frequency (VHF) and ultra high frequency (UHF), a vehicle moving at 50km/hr

or 14 m/sec, will pass through several fades in a second [11]. Deep fades can also be caused by the continuous physical changes of the channel media. When the interference and fluctuation are much more rapid than the signaling rate, fast fading occurs.

Frequently, fading and multipath phenomena occur simultaneously. However, there are situations, such as stationary communications, in which both the transmitter and the receiver are not moving, where the main cause of degradation of communication quality is the occurrence of multipath transmission. There are also situations like communication between high-speed vehicles in an environment with no major reflective surfaces, where the main cause of degradation is fading. To make the multipath-fading channel problem more tractable, it is helpful to consider the multipath aspect or the fading aspect separately, and this is the approach we use.

1.2 Literature Review

The history of fading channel studies can be roughly divided into three periods. The first period spans from the early 1950's to the late 1960's. The driving force behind this research was mainly the impetus for the development of long-distance troposcatter communication for military use [14,15,25–27]. Tremendous effort was invested in collecting real data to illustrate fading channels [28]. Some practical mathematical models, such as uncorrelated scattering, were introduced [14, 17]. The single most important idea for combating fading phenomenon, "diversity," was also formulated in this period [29–36]. The beneficial effect of diversity vanishes when the channel has a very high correlation, i.e., no fading [22]. A lot of work was also performed in the area of detection. However, systems were generally not reliable enough for some computer communication applications. This was the main reason why coding theory was introduced [19].

The second period spans from the early 1970's to the mid 1980's. During this period, there was a drop in interest in fading channel communication, which might be due to the high demand in research of wire and optical communications, where fading is not an important issue. Even so, there was still continuous effort in the information and coding theory aspects of fading channel research [37–40].

The third period spans from the mid 1980's to the present. During this period, there has been a huge increase in scientific activity in fading channel research. Many new results have been reported. The main driving force is the tremendous increase in demand for personal mobile wireless communication, of which the following references specify aspects of the research. The research in this period can be roughly divided into three categories, i.e., implementation, characterization, and optimization.

1. Implementation:

   Many new implementation schemes have been derived and reported. These approaches can be further subdivided into three more classes: a) Diversity, b) Detection and Estimation, and c) Combined Methods.

   (a) Diversity: The application of the idea of diversity exists in virtually every branch of research on fading channels, either implicitly or explicitly. A basic motivation for diversity is that under fairly general conditions, a channel affected by fading can be transformed into an additive-white-Gaussian-noise (AWGN) channel by increasing the branches of diversity [36,41–45]. There are mainly three types of diversity, i.e., time, space, and polarization diversity. Time diversity is usually achieved by using some kind of rearrangement of the signaling process [46], or by means of coding [35, 41, 47, 48]. Space diversity can be achieved by using multiple antennas at the transmitter and/or the receiver. Polarization diversity is achieved by using an antenna or antennas with multiple polarizations [49–52].

   (b) Estimation and Detection: There are hybrid adoptions of classic techniques of estimation and detection theory for fading situations, usually under specific settings [53–56]. For example, there are applications of diversity on the design of detector and estimator [36,57], and the usage of robust adaptive algorithms in equalizer designs for fast fading channels [58–66].

   (c) Combined Methods: These methods include information theory and spatial diversity [67], adaptive code-division-multiple-access (CDMA) signaling [68], space-time modulation [49], and the applications of channel coding [69,70].

2. Characterization:

   Tremendous effort has been devoted to the study of channel characterization, the polishing of necessary mathematical tools, and the development of performance evaluation techniques. Furthermore, theoretical performance bounds have been derived, especially for matched filter detectors.

   (a) Channel Characterization: Channel characterization is generally carried out on an experimental basis. Measurement of transmitted and received signals are made in typical environments in a controlled manner. Statistical models are formulated and validated against the measured data and applied to system analysis and simulation. Some researchers collect their own data and construct their own channel models [71] while others study the existing experimental data with their own models or improve previous models [72, 73]. The demand for research in this topic has also inspired the development in specific mathematical techniques [46, 74–76] and new statistical models [77]. The characteristics and effects of the channel on the various aspects of system performance (e.g., error rate under specific settings, channel capacity, etc.) alone have also induced much interest among researchers [78–81].

   (b) Simulation Techniques: In the design of a communication system, a system designer may want to ensure that the performance of the communication link is satisfactory. To ensure the system is performing up to the design specifications, a computer or hardware simulation is required. Therefore, many different simulation schemes have been proposed for different settings [76, 82, 83].

   (c) Performance Analyses: Performance analyses have been derived for uses in nearly all common fading situations and standard communication techniques. These analyses are often case-specific [42, 43, 45, 46, 51, 74, 84–92].

   (d) Performance Bounds: Different performance bounds that have been derived now set guidelines to obtain optimal performance for different fading channels

[93,94]. Among different performance bounds, the matched-filter-bound (MFB) has received special interest from researchers [34,95–99]. The MFB characterizes the detection capability of a single communication pulse transmitted in isolation so that potential effects of ISI can be negligible. Although the MFB is a lower bound over a communication channel that may not be practically realizable, it is a common performance measurement for communication systems [96, 97]. There are reasons for saying MFB is an optimal performance bound. First, it is assumed that all the information about the channel is known, which is not feasible for a real fading channel. Second, transmitted pulses are separated sufficiently so that no ISI occurs. In other words, the error rate is determined by assuming only one data pulse is transmitted, which is also not physically realistic.

3. Optimization:

   A great amount of valuable work has gone into assessing the information-theoretic limits of fading channels. Tremendous effort has been devoted to the derivation of channel capacity and the search for optimal signals to achieve its optimal capacity.

   (a) Channel Capacity: The capacity with or without the knowledge of channel state information has been studied [67,100,101], or with some further constraints, such as the delay limited to real time voice and video communication [102]. Goldsmith [100] and Caire [101] interpreted the optimal power adaptation as "water-pouring" in time or frequency, while for multipath-fading channels, Telatar [103] interpreted these as "peaky" in time or frequency. For multi-access usage of the fading channel, there is a solution analogous to the "water-pouring" interpretation for the single-user case to maximize the overall capacity [104]. In general, the capacity as well as the capacity-achieving distribution implies that there are some underlying structures of optimal coding/signaling [105].

   (b) Optimal Power Allocation: With the assumption of block-fading additive-white-Gaussian-noise (BF-AWGN) and the knowledge of channel states, an optimal

power control for both the transmitter and the receiver is derived with or without the presence of a transmission-delay constraint [106, 107]. A discrete observable, usually orthogonal projections of the tested signal onto a family of specific functions, has been used to reduce the receiver complexity in the study of the performance for difference classes of signaling [108]. Also, there are more general results for signaling properties for fading channels [109, 110].

1.3 Methodology

The research in this dissertation utilizes the idea of separating the multipath aspect from the fading aspect of a fading channel in order to reduce the model complexity. The goal of this dissertation is to characterize the optimal signal and detector of a fading channel, which may be either fast fading or slow fading, when its channel statistical properties are known.

The word "optimal" used in this dissertation denotes maximized signal-to-noise-ratio (SNR). Improved SNR is related to but not necessarily equivalent to improved error rate [92]. Optimization of SNR does not necessarily mean optimization of the error rate. Optimizing in terms of error rate is the ultimate goal for all communicators. It involves sophisticated modeling of the entire communication system including the optimal detector corresponding to the autocorrelation function (which is a stochastic process itself) of the fading channel. At the present stage of research, we do not try to tackle this problem, taking on instead the more tractable optimization of SNR. This is analogous to the SNR-maximizing derivation of the conventional matched filter, which also happens to result in minimal bit error rate in binary-phase-shift-keying (BPSK) systems. In any event, improvement in received SNR will result in an improvement in the bit error rate.

We first characterize optimal detectors, then study optimal signaling. Because of our model setting, these two optimization problems are essentially the same mathematically, which significantly reduces the difficulty of the derivation and understanding of the characterization of the optimal-signal-filter-pair for fading channels. For the sake of clarity of later discussion, we call this optimal-signal-filter-pair a **consonant pair** (CP).

We use functional analysis techniques to derive the characteristics of the optimal detector and signal in $\mathcal{L}_2$.[1] We then further analyze the properties of the optimal filters and signals, and present some algorithms that help us to find these properties.

Without the assumption of slow fading, the matched filter is no longer necessarily the optimal detector [111, 112]. As a result, the SNR of the received signal will depend strongly on the channel statistical properties, the corresponding transmitted function, and the detection filter. Therefore, with proper functional analysis techniques, we can derive some basic properties of the signal and filter pairs to maximize the SNR.

There are two main results from this research. First, maximizing the SNR of the detected signal is a key factor to lowering the error rate in communication, which implies a higher possible data rate, increased channel re-usability, smaller appliance, etc. Second, a better characterization of the optimal signal enables us to have better resource management, e.g., if we know that the support of the optimal signal is related to the support of the autocorrelation function of the channel, we will not slow down our transmission rate by using long signals.

1.4 Contributions to the Field of Knowledge

After an intensive literature search, we believe the following contributions to the field of optimal detector and signaling in fading channels are original to this dissertation:

1. A characterization of the optimal detector for known signals in fading channels. We find that the well known matched filter and RAKE receiver are limiting cases of our result.

2. A characterization of the CP in terms of maximum SNR for given fading channels. Not only is fading (multiplicative and multipath) considered as a channel effect, but also the effect of synchronization uncertainty on the optimal signal/detector design.

[1] $\mathcal{L}_2(\mathbb{R})$ is the collection of all Lebesgue measurable functions $f : \mathbb{R} \to \mathbb{C}$, for which $|f|^2$ is Lebesgue integrable, i.e., $\int_{-\infty}^{\infty} |f|^2 < \infty$. In this dissertation, we use $\mathcal{L}_2$ to represent $\mathcal{L}_2(\mathbb{R})$.

3. An iterative algorithm for finding CPs that can be used for communication through fading channels. With land-mobile-fading channels, the signals obtained with our procedure are demonstrated to have a performance significantly better than that of the traditional flat-top pulse and raised cosine pulse in a PPM pair in terms of probability of error rate at a comparable signaling rate.

4. A closed-form expression for the probability density function and the cumulative distribution function of quadratic form, $g = \mathbf{v}^H Q\mathbf{v}$, where $\mathbf{v}$ is a complex Gaussian vector with $E(\mathbf{v}) = 0$, and Q is a Hermitian matrix. The probability density function and the cumulative distribution function are expressed in terms of the eigenvalues of MQ, where M is the covariance matrix of $\mathbf{v}$. The closed-form expression is applicable to many general cases. In this dissertation, we apply the equation to the calculation of the probability of error rate for quadratic receivers. As a corollary to this analysis, an interesting partition of unity by sequences is presented. From the above cumulative distribution function of the quadratic form of normal vectors, we are also able to derive an expression of partition of unity which is shown to be very useful.

The results of this dissertation are primarily analytical, with the structure of the CPs being developed. Nevertheless, the concepts derived herein have been implemented, and a description of this implementation is provided. Also a modest set of computed results is presented to demonstrate the nature of the solutions and the algorithms. While an exhibitive demonstration is not provided, the tools described here can be used in a variety of circumstances.

1.5 Dissertation Structure

In chapter 2, we illustrate the fading channel model that is commonly used. Following this, we introduce the model which we have used in our research. In chapter 3, we discuss the optimal detection of a given signal in fading channels. Then, we relate the optimal signal to its optimal detection. In chapter 4, we develop some algorithms that we can use to generate CPs. As we only characterize a single CP, while a realistic communication requires

at least two different signals to represent "0" and "1", in the last section of chapter 4, we discuss a procedure to find two conjugate-consonant pairs which can be used for binary transmission. In chapter 5, we derive a method to calculate the probability of error for our channel model and compare the CPs obtained in our research to other traditional signals. Next, we discuss the issues of implementing our algorithms with discrete samples in chapter 6. Results of the implementation are presented in chapter 7 along with observations and discussions. We conclude this dissertation in chapter 8 with a list of topics for further research.

Chapter 2
Channel Models

2.1 Introduction

Fading channels were first modeled in the 1950s and 1960s as a large number of "scatterers" located at random points within the propagation path [16,18]. This idea was primarily applied to over-the-horizon (troposcatter) communications covering a wide range of frequency bands. To reach beyond the horizon, clouds of particles in the troposphere were used as reflectors for the radio waves. The differences in path length between the large number of scattered waves give rise to Rayleigh fading if there is no dominant direct component [27]. Mobile wireless systems of recent interest, which are mainly local and with finite number of reflection paths, experience fading effects that are somewhat different than those mentioned above. However, these early models are still quite useful to help characterize fading effects that we are facing nowadays.

The groundwork of modeling fading channels was mathematically laid out by Price in the mid 1950's [14,15]. For the following half century, there were tremendous contributions to this field. In a recent survey paper written by Biglieri, Proakis, and Shamai, 549 citations reporting the state-of-the-art achievements in research on fading channels were included [22]. Among those that are not included in their citations is a paper written by Bello who introduced a simple way to model the fading phenomenon with the notion of wide-sense stationary uncorrelated scattering (WSSUS) [17]. Uncorrelated scattering means the attenuation and phase shift of the channel associated with different path delays are not correlated.

2.2 Fading Channel Model

A widely accepted mathematical modeling of the fading channel can be found in Proakis's book [21]. A fading channel is viewed as a continuous time-varying filter with a

baseband-equivalent impulse response of $c(\tau;t)$, where $c(.;.)$ is a complex valued function, τ is the delay in response and t is the time variable. If a signal $s(t)$ is sent through this channel with AWGN, we will have at the receiver side,[1]

$$r(t) = \int c(\tau;t)s(t-\tau)d\tau + n(t), \tag{2.1}$$

where $r(t)$ is the received signal and $n(t)$ is the AGWN.

Assuming $c(\tau;t)$ is wide-sense stationary, the autocorrelation function of $c(\tau;t)$ can be defined as

$$\phi_c(\tau_1,\tau_2;\Delta t) = \frac{1}{2}E[c^*(\tau_1;t)c(\tau_2;t+\Delta t)]. \tag{2.2}$$

With uncorrelated scattering, we have

$$\phi_c(\tau_1;\Delta t)\delta(\tau_1-\tau_2) = \frac{1}{2}E[c^*(\tau_1;t)c(\tau_2;t+\Delta t)]. \tag{2.3}$$

Setting $\Delta t = 0$, the expression $\phi_C(\Delta\tau;0)$ in (2.3) is usually called the **Multipath Intensity profile** or the **Delay Power Spectrum** of the fading channel. We further define T_m to be the length of time that $\phi_C(\Delta\tau;0)$ is essentially nonzero, and call it the **Multipath Spread** of the fading channel.

Taking the Fourier transform of the channel impulse response, we have

$$C(f;t) = \int c(\tau;t)e^{-j2\pi f\tau}d\tau. \tag{2.4}$$

The corresponding autocorrelation function, with WSSUS, is

$$\begin{aligned}
\phi_C(f_1,f_2;\Delta t) &= \frac{1}{2}E[C^*(f_1;t)C(f_2;t+\Delta t)] \\
&= \frac{1}{2}\int\int E[c^*(\tau_1;t)c(\tau_2;t+\Delta t)]e^{j2\pi(f_1\tau_1-f_2\tau_2)}d\tau_1 d\tau_2 \\
&= \int\int \phi_c(\tau_1;\Delta t)\delta(\tau_1-\tau_2)e^{j2\pi(f_1\tau_1-f_2\tau_2)}d\tau_1 d\tau_2 \\
&= \int \phi_c(\tau_1;\Delta t)e^{j2\pi(f_1-f_2)\tau_1}d\tau_1 \\
&= \int \phi_c(\tau_1;\Delta t)e^{-j2\pi\Delta f\tau_1}d\tau_1 \\
&\equiv \phi_C(\Delta f;\Delta t)
\end{aligned} \tag{2.5}$$

[1]Throughout this work, integrals stated without limits are assumed to be over the interval $(-\infty,\infty)$.

Setting $\Delta t = 0$, we obtain $\phi_C(\Delta f; 0)$; this is the autocorrelation function in the frequency variable. We define the length of the support of $\phi_C(\Delta f; 0)$ that is essentially nonzero as $(\Delta f)_C$. Usually, $(\Delta f)_C$ is denoted as the **Coherence Bandwidth**, which is a measure of the frequency coherence of the fading channel. To illustrate this idea, we would imagine a signal with a bandwidth less than $(\Delta f)_C$ being transmitted through a fading channel. Because all frequency components will be affected essentially the same by the channel, this channel is said to be frequency-nonselective. However, if another signal with a bandwidth larger than $(\Delta f)_C$ is being transmitted through the same fading channel, different frequency components will be affected in different ways by the channel. Then, the channel is said to be frequency-selective. In this case, the signal can be severely distorted by the channel. Another useful result due to the fact that $\phi_C(\Delta\tau; 0)$ and $\phi_C(\Delta f; 0)$ are Fourier transform pairs is that

$$(\Delta f)_C \approx \frac{1}{T_m}. \tag{2.6}$$

To study the time variation of the channel, we return to $\phi_C(\Delta f; \Delta t)$ defined in (2.5). By setting $\Delta f = 0$, we obtain the autocorrelation function of the channel for each different frequency component by time averaging. Now, the length of the support of $\phi_C(0; \Delta t)$ that is essentially nonzero is defined as $(\Delta t)_C$, which is called the **Coherence Time** of the fading channel. So, in a digital communication system, if the signaling period is shorter than $(\Delta t)_C$, we may assume that the channel is essentially constant for individual signals, and the situation of slow fading occurs. On the other hand, if the signaling period is longer than $(\Delta t)_C$, we may no longer assume that the channel is constant for individual signals, and fast fading occurs.

In practice, $(\Delta t)_C$ is very difficult to measure directly. So, a related function is employed. This new function is called the scattering function of the fading channel. It is defined as

$$S(\tau; \zeta) = \int\int \phi_C(\Delta f; \Delta t) e^{-2\pi\tau\Delta f} e^{-2\pi\zeta\Delta t}\, d\Delta f\, d\Delta t, \tag{2.7}$$

where τ is the time delay, and ζ is the Doppler frequency as the channel varies with time (caused by physical motions of the media, the transmitter, and the receiver, etc). The

scattering function of the channel provides us with a measure of the average power output of the channel. When $\tau = 0$, the length of the support of $S(0;\zeta)$ that is essentially nonzero is called the **Doppler spread** B_d of the fading channel. We have the following relationship between the channel coherence time and the Doppler spread

$$(\Delta t)_C \approx \frac{1}{B_d}. \tag{2.8}$$

The time variation autocorrelation function, $\phi_C(0;\Delta t)$, can be obtained from $S_C(\zeta) \equiv S(0;\zeta)$, and is called the **Doppler power spectrum** of the fading channel. Some of the common spectra and autocorrelation functions are listed in Table 2.1 [9, 84, 92, 112–114].

Table 2.1: Spectra and Autocorrelation Function of Fading Process.

Denotation	Spectrum $S_C(\zeta)$	Autocorrelation Function $R_C(\tau)$
1. Rectangular	$\frac{1}{2B_d}$ $\quad \lvert\zeta\rvert < B_d$	$K\frac{\sin 2\pi B_d\tau}{2\pi B_d\tau}$
2. Gaussian	$K\exp(\frac{-\zeta^2}{{B_d}^2})/\sqrt{\pi}B_d$	$K\exp[-(\pi B_d\tau)^2]$
3. Land Mobile	$\frac{K}{\pi(\zeta^2 - {B_d}^2)^{1/2}}$	$KJ_0(2\pi B_d\tau)$
4. First-Order Butterworth	$\frac{K}{\pi B_d(1+\zeta^2/{B_d}^2)}$	$K\exp[-2\pi B_d\lvert\tau\rvert]$
5. Second-Order Butterworth	$\frac{K}{1+16\frac{\zeta^4}{{B_d}^4}}$	$K\exp[-\frac{\pi B_d\lvert\tau\rvert}{\sqrt{2}}]$ $\times\left(\cos\frac{\pi B_d\lvert\tau\rvert}{\sqrt{2}} + \sin\frac{\pi B_d\lvert\tau\rvert}{\sqrt{2}}\right)$

Note: $J_0(\cdot)$ is the zero-order Bessel function of the first kind, and K is a constant.

2.3 Categories of Degradation

Different fading phenomena introduce different degradations in communication systems. These degradations are summarized in Table 2.2 [23].

Table 2.2: Categories of Degradation According to Different Fading Phenomena.

Frequency-nonselective	Frequency-selective	Slow fading	Fast fading
loss in SNR	ISI distortion, pulse dispersion, irreducible BER	low Doppler, loss in SNR	high Doppler, PLL failure, irreducible BER

2.4 Model for Frequency-Nonselective Fading Channels

Let us revisit (2.1), i.e.,

$$r(t) = \int c(\tau; t)s(t-\tau)d\tau + n(t).$$

By Parseval's relation, we have

$$r(t) = \int C(f; t)S(f)e^{j2\pi ft}df + n(t). \tag{2.9}$$

Assuming frequency nonselectiveness, we have $C(f;t) = C(0;t)$, and

$$\begin{aligned} r(t) &= \int C(f; t)S(f)e^{j2\pi ft}df + n(t) \\ &= C(0;t)\int S(f)e^{j2\pi ft}df + n(t) \\ &= C(0;t)s(t) + n(t). \end{aligned} \tag{2.10}$$

The transfer function $C(0;t)$ for a frequency-nonselective fading channel may be expressed as

$$C(0;t) = \alpha(t) = a(t)e^{-j\theta(t)}, \tag{2.11}$$

where the random process $a(t)$ represents the envelope, and the random process $\theta(t)$ represents the phase of the the transfer function. In this case, the channel is also called a multiplicative fading channel. Note that a multiplicative channel can be both fast- and slow-fading. For different fading channels, $a(t)$ can acquire different statistical distributions, from which each different fading channels obtains its name. Among common distributions for $a(t)$ are the Rayleigh distribution, the Rician distribution, and the Nakagami distributions [92].

2.5 Model Used in This Dissertation

The approach that is employed in this dissertation for modeling the general fading channel is first to assume frequency-nonselectiveness, then to introduce countably many multiple paths when we want to study the frequency-selectiveness aspect of the fading channel. This approach agrees with the assumption of independent fading paths put forward by Bello [17]. Furthermore, this approach provides a model that is general enough to represent most "real" fading channels that are used in practice. At the same time, the model also allows us to separate the fading and multipath aspects of a fading channel and deal with them separately.

Fig. 2.1 depicts the system that we are studying. It is a generic communication system for fading channels with a correlation detector.

When we assume frequency-nonselectiveness, a fading channel becomes simply a multiplicative Gaussian channel, i.e., when $s(t)$ is sent, we receive

$$r(t) = \alpha(t)s(t) + n(t). \tag{2.12}$$

Without loss of generality, we may assume $\alpha(t)$ has a unit variance, i.e., $E[\alpha(t)\alpha^*(t)] - E[\alpha(t)]E[\alpha^*(t)] = 1$. For a general multipath fading channel, we receive

$$r(t) = \sum_i \xi_i \alpha_i(t) s(t - \phi_i) + n(t), \tag{2.13}$$

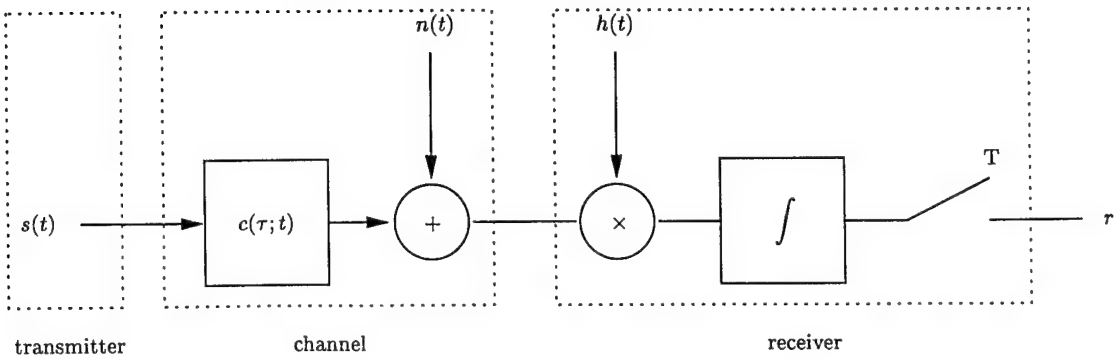


Fig. 2.1: System model of a fading channel.

where $\alpha_i(t)$ has an unit variance, ξ_i is the path gain for path i, and ϕ_i is its path delay. The autocorrelation functions of $\alpha_i(t) = p_i(t) + jq_i(t)$, where both $p_i(t)$ and $q_i(t)$ are real random processes, is

$$\begin{aligned} R_{\alpha_i}(\tau,\lambda) =& E[\alpha_i(\tau)\alpha_i^*(\lambda)] \\ =& E[(p_i(\tau)p_i(\lambda) + q_i(\tau)q_i(\lambda)) - j(p_i(\tau)q_i(\lambda) - q_i(\tau)p_i(\lambda))] \\ =& E[(p_i(\tau)p_i(\lambda)] + E[q_i(\tau)q_i(\lambda)] \end{aligned} \tag{2.14}$$

where it is assumed that the real and imaginary parts are independent. The autocorrelation $R_{\alpha_i}(\tau,\lambda)$ is real and symmetric about τ and λ.

When $\alpha_i(t)$ is Rayleigh, i.e., $p_i(t)$ and $q_i(t)$ are zero mean Gaussian processes, and when we further consider that the different fading paths are only caused by local reflections, it is safe to say that the $\alpha_i(t)$ all have the same normalized statistical characteristics, i.e., we have for some autocorrelation function $R_\alpha(\tau,\lambda)$ that $R_{\alpha_i}(\tau,\lambda) = R_\alpha(\tau,\lambda)$ for all i. The above simplification cannot be applied in general to other fading channels. For example, if $\alpha_i(t)$ is Rician distributed, the directions of the dominant components are different for each of the multiple paths, and hence the autocorrelation functions may be different. To simplify the derivation, we will assume $\alpha(t)$ to be Rayleigh in this dissertation. However, our discussion is readily extendible to other fading channels.

If we do not assume perfect synchronization when a known signal $s_0(t)$ is transmitted, the received signal at the output of the detector filter is

$$r = \int [\sum_i \xi_i \alpha_i(t) s_0(t - \phi_i - \mu) + n(t)] h(t)\, dt,$$

where μ is a random delay with density function $p(\mu)$. We assume throughout this work that $h(t)$ has unit energy, i.e., $\int h(t)h^*(t)\, dt = 1$.

Chapter 3
Optimal Detection and Signaling

Up to the present, research about fading channels has been performed mostly under the slow fading assumption, i.e., the channel is considered to be constant for up to a few symbol times. Under this assumption, a fading channel behaves as an AWGN channel for each symbol. It is also generally accepted that matched filter detectors produce good performance. However, when the slow fading assumption is lifted, the idea of optimal detection becomes more complicated. Fast fading causes variations within a symbol interval, so that the model applied in slow fading becomes inaccurate, and conventional matched filter detectors become suboptimal. In this chapter, we characterize optimal detection and signaling without the assumption of slow fading.

3.1 Optimal Detection

For a single-path frequency-nonselective fading channel with random amplitude function, $\alpha(t)$, when a known signal $s_0(t)$ is transmitted in a Gaussian channel, the received signal is

$$r(t) = \alpha(t)s_0(t-\mu) + n(t), \tag{3.1}$$

where μ is a random delay with probability density function $p(\mu)$ due to uncertainty in the synchronization, and $n(t)$ is AWGN with a power-spectral-density (PSD) $N_0/2$. For a multipath fading channel, the received signal is

$$r(t) = \sum_i \xi_i \alpha_i(t) s_0(t-\phi_i-\mu) + n(t), \tag{3.2}$$

where ξ_i is the channel gain, and ϕ_i is the path delay for path i. We would like to find the optimal $h(t)$, with unit energy $\int h(t)h^*(t)\,dt = 1$, that maximizes the received SNR.

3.2 Characterization of Optimal Detectors

The received signal at the output of a linear detection filter can be presented as

$$r = \int [\sum_i \xi_i \alpha_i(t) s_0(t - \phi_i - \mu) + n(t)] h(t)\, dt, \tag{3.3}$$

with the corresponding signal and noise components at the output of the detection filter being

$$r_s = \int \sum_i \xi_i \alpha_i(t) s_0(t - \phi_i - \mu) h(t)\, dt, \quad \text{and} \quad r_n = \int n(t) h(t)\, dt. \tag{3.4}$$

Since $h(t)$ has unit energy, we have

$$E(|r_n|^2) = \frac{N_0}{2}. \tag{3.5}$$

The variance for the signal part is

$$\begin{aligned} E(|r_s|^2) =& E[\int\int (\sum_i \xi_i \alpha_i(\tau) s_0(\tau - \phi_i - \mu)) h(\tau) (\sum_j \xi_j^* \alpha_j^*(\lambda) s_0^*(\lambda - \phi_j - \mu)) h^*(\lambda)\, d\tau\, d\lambda] \\ =& \sum_i \sum_j \int\int \xi_i \xi_j^* E[\alpha_i(\tau) \alpha_j^*(\lambda) s_0(\tau - \phi_i - \mu) s_0^*(\lambda - \phi_j - \mu)] h(\tau) h^*(\lambda)\, d\tau\, d\lambda. \end{aligned} \tag{3.6}$$

Since $\alpha(t)$ and $s_0(t)$ are independent, we have

$$E(|r_s|^2) = \sum_i \sum_j \int\int \xi_i \xi_j^* E[\alpha_i(\tau) \alpha_j^*(\lambda)] E[s_0(\tau - \phi_i - \mu) s_0^*(\lambda - \phi_j - \mu)] h(\tau) h^*(\lambda)\, d\tau\, d\lambda. \tag{3.7}$$

Because of the assumption that different paths are independent, we obtain

$$E(|r_s|^2) = \sum_i \int\int |\xi_i|^2 E[\alpha_i(\tau) \alpha_i^*(\lambda)] E[s_0(\tau - \phi_i - \mu) s_0^*(\lambda - \phi_i - \mu)] h(\tau) h^*(\lambda)\, d\tau\, d\lambda. \tag{3.8}$$

Defining

$$\begin{aligned} \Psi_i(\tau, \lambda) &= E[s_0(\tau - \phi_i - \mu) s_0^*(\lambda - \phi_i - \mu)] \\ &= \int s_0(\tau - \phi_i - \mu) s_0^*(\lambda - \phi_i - \mu) p(\mu)\, d\mu, \end{aligned} \tag{3.9}$$

(3.8) can be rewritten as

$$\begin{aligned} E(|r_s|^2) =& \sum_i \int\int |\xi_i|^2 R_\alpha(\tau, \lambda) \Psi_i(\tau, \lambda) h(\tau) h^*(\lambda)\, d\tau\, d\lambda \\ =& \int\int R_\alpha(\tau, \lambda) \{\sum_i |\xi_i|^2 \Psi_i(\tau, \lambda)\} h(\tau) h^*(\lambda)\, d\tau\, d\lambda. \end{aligned} \tag{3.10}$$

Let

$$\Psi(\tau,\lambda)=\sum_i |\xi_i|^2\Psi_i(\tau,\lambda). \tag{3.11}$$

Then, the variance for the signal part may be presented as a functional of h.

$$\begin{aligned} E(|r_s|^2) &= \iint R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)h(\tau)h^*(\lambda)\,d\tau\,d\lambda \\ &\stackrel{\text{def}}{=} F(h). \end{aligned} \tag{3.12}$$

Note that both $\Psi(\tau,\lambda)$ and $R_\alpha(\tau,\lambda)$ are symmetric in τ and λ.

By applying the method of Lagrange multipliers to maximize $F(h)$ under the constraint of $H(h)=\int h(t)h^*(t)\,dt=1$, we have the following Lagrangian function

$$J(h)=F(h)-\gamma H(h), \tag{3.13}$$

where γ is the Lagrange multiplier. To find the extrema, we take the derivative with respect to h (see Appendix A) of the Lagrangian and set it to zero, i.e.,

$$\begin{aligned} &\frac{d}{dh}J(h_0)=\frac{d}{dh}F(h_0)-\gamma\frac{d}{dh}H(h_0)=0 \\ \Rightarrow &\frac{d}{dh}F(h_0)=\gamma\frac{d}{dh}H(h_0), \end{aligned} \tag{3.14}$$

where $h_0(t)$ is an extremum of $F(h)$. From the example in Appendix A, we know that $\frac{dH(h_0)}{dh}(\tau)=h_0^*(\tau)$. Now, we try to find $\frac{dF(h_0)}{dh}(\tau)$. For $\epsilon\in\mathbb{R}$, and $V\in\mathcal{L}_2$, and $h_0(t)=u_0(t)+jv_0(t)$,

$$\begin{aligned} &F((u_0+\epsilon V)+jv_0)-F(h_0) \\ &=\iint R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)((u_0+\epsilon V)+jv_0)(\tau)((u_0+\epsilon V)-jv_0)(\lambda)d\tau d\lambda-F(h_0) \\ &=\iint R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)\{(u_0+\epsilon V)(\tau)(u_0+\epsilon V)(\lambda)-j(u_0+\epsilon V)(\tau)v_0(\lambda) \\ &\quad +jv_0(\tau)(u_0+\epsilon V)(\lambda)+v_0(\tau)v_0(\lambda)\}\,d\tau\,d\lambda-F(h_0). \end{aligned} \tag{3.15}$$

By neglecting those terms containing ϵ^2, (3.15) becomes

$$\begin{aligned} &F((u_0+\epsilon V)+jv_0)-F(h_0) \\ &\approx\iint R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)\{u_0(\tau)u_0(\lambda)+\epsilon V(\tau)u_0(\lambda)+\epsilon u_0(\tau)V(\lambda)-ju_0(\tau)v_0(\lambda) \\ &\quad -j\epsilon V(\tau)v_0(\lambda)+jv_0(\tau)u_0(\lambda)+j\epsilon v_0(\tau)V(\lambda)+v_0(\tau)v_0(\lambda)\}\,d\tau\,d\lambda-F(h_0) \\ &=\iint R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)\{\epsilon V(\tau)u_0(\lambda)+\epsilon u_0(\tau)V(\lambda)-j\epsilon V(\tau)v_0(\lambda)+j\epsilon v_0(\tau)V(\lambda)\}\,d\tau\,d\lambda. \end{aligned} \tag{3.16}$$

As both R_α and Ψ are symmetric in τ and λ, (3.16) may be further simplified as

$$\begin{aligned} & F((u_0+\epsilon V)+jv_0)-F(h_0) \\ =&\epsilon\int\int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)\{u_0(\lambda)+u_0(\lambda)-jv_0(\lambda)+jv_0(\lambda)\}V(\tau)\,d\tau\,d\lambda \\ =&2\epsilon\int\int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)u_0(\lambda)V(\tau)\,d\tau\,d\lambda. \end{aligned} \tag{3.17}$$

From Appendix A, we have

$$\frac{\partial F(h_0)}{\partial u}(\tau)=2\int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)u_0(\lambda)\,d\lambda. \tag{3.18}$$

Similarly,

$$\frac{\partial F(h_0)}{\partial v}(\tau)=2\int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)v_0(\lambda)\,d\lambda. \tag{3.19}$$

From (3.18) and (3.19), we have

$$\begin{aligned} \frac{d\,F(h_0)}{dh}(\tau) =&\frac{1}{2}\left(\frac{\partial\,F(h_0)}{\partial\,u}(\tau)-j\frac{\partial\,F(h_0)}{\partial\,v}(\tau)\right) \\ =&\int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)h_0^*(\lambda)\,d\lambda. \end{aligned} \tag{3.20}$$

So, from (3.14),

$$\gamma h_0^*(\tau)=\int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)h_0^*(\lambda)\,d\lambda. \tag{3.21}$$

Also,

$$\begin{aligned} E(|r_s|^2) =&F(h_0)=\int\int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)h_0(\tau)h_0^*(\lambda)\,d\tau\,d\lambda \\ =&\int \gamma h_0(\tau)h_0^*(\tau)\,d\tau=\gamma. \end{aligned} \tag{3.22}$$

We can conclude that in order to maximize the SNR, γ must be the largest eigenvalue of the kernel $R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)$, and $h_0^*(t)$ is the corresponding eigenfunction.

Assuming perfect synchronization, i.e., $p(\mu)=\delta(\mu)$, it is interesting to point out that for the case of a single nonfading path, i.e., $R_\alpha(\tau,\lambda)=R$ (constant) for all τ and λ,

$$\begin{aligned} \gamma h_0^*(\tau) =&\int R_\alpha(\tau,\lambda)s_0(\tau)s_0^*(\lambda)h_0(\lambda)\,d\lambda \\ =&s_0(\tau)\int Rs_0^*(\lambda)h_0^*(\lambda)\,d\lambda \\ =&Cs_0(\tau), \end{aligned} \tag{3.23}$$

where $C\in\mathbb{C}$. This is the familiar matched filter.

On the other hand, for the case of nonfading multiple paths,

$$\begin{aligned}\gamma h_0^*(\tau) &= \sum_i \{|\xi_i|^2 \int R s_0^*(\lambda - \phi_i)) h_0(\lambda)\, d\lambda\} s_0(\tau - \phi_i) \\ &= \sum_i C_i s_0(\tau - \phi_i),\end{aligned} \tag{3.24}$$

which is the RAKE receiver [21]. The derivation presented above subsumes and generalizes these important special cases.

3.3 Optimal Signaling

We would like to extend our discussion now to the study of how to characterize both the transmitted signal $s(t)$ and the detection filter $h(t)$ so that for a given fading channel with autocorrelation function $R_\alpha(\tau, \lambda) = E(\alpha(\tau)\alpha^*(\lambda))$, the received SNR at the receiver is maximized.

For a fading channel, $c(\tau; t)$ whose statistical properties are known, when the signal $s(t)$ is sent, the received waveform can be represented as

$$r(t) = c(\tau; t) * s(t - \tau) + n(t), \tag{3.25}$$

where $n(t)$ is an AWGN with power spectral density $N_0/2$. Consider the problem of designing a filter $h(t)$ matched to $c(\tau; t) * s(t)$, then shaping $s(t)$ to obtain the maximum SNR possible. That is, find

$$s_{\max}(t) = \arg \max_{\substack{s(t) \\ \int |s(t)|^2 dt = 1}} \left\{ \max_{\substack{h(t) \\ \int |h(t)|^2 dt = 1}} \frac{\left| \int [c(\tau; t) * s(t)] h(t) dt \right|^2}{N_0} \right\}. \tag{3.26}$$

3.4 Characterization of Optimal Signaling for Single-Path Fading

We derive the CP first for single path fading with no synchronization error. This will bring us better understanding of the properties of this optimization problem (3.26). We would like to characterize the optimal $s(t)$ and $h(t)$ so that for a given $R_\alpha(\tau, \lambda)$, the SNR at the receiver is maximized. We impose the constraint that both $s(t)$ and $h(t)$ have unit energy,

$$G(s) = \int s(t) s^*(t) dt = 1 \quad \text{and} \quad H(h) = \int h(t) h^*(t) dt = 1. \tag{3.27}$$

Without synchronization uncertainty, the received signal at the output of a detector filter is

$$r = \int [\alpha(t)s(t) + n(t)]h(t)\ dt, \tag{3.28}$$

with the corresponding signal and noise components being

$$r_s = \int \alpha(t)s(t)h(t)\ dt, \quad \text{and} \quad r_n = \int n(t)h(t)\ dt. \tag{3.29}$$

Maximizing the SNR is equivalent to maximizing the ratio between the variance of r_s and r_n. Because of (3.27), we have $E(|r_n|^2) = N_0/2$. So, what we really need to do is to maximize the variance of r_s, which can be represented as the following functional

$$\begin{aligned} E(|r_s|^2) =& E[\int\int \alpha(\tau)s(\tau)h(\tau)\alpha^*(\lambda)s^*(\lambda)h^*(\lambda)d\tau d\lambda] \\ =& \int\int E[\alpha(\tau)\alpha^*(\lambda)]s(\tau)h(\tau)s^*(\lambda)h^*(\lambda)d\tau d\lambda \\ =& \int\int R_\alpha(\tau,\lambda)s(\tau)h(\tau)s^*(\lambda)h^*(\lambda)d\tau d\lambda \\ \stackrel{\text{def}}{=}& F(s,h). \end{aligned} \tag{3.30}$$

Using the method of Lagrange multipliers to maximize $F(s,h)$, i.e., we construct the Lagrangian

$$J(s,h) = F(s,h) - \gamma_0 G(s) - \gamma_1 H(h), \tag{3.31}$$

and set the gradient of $J(s,h)$ at (s_0,h_0) to be zero, where $s_0, h_0 \in \mathcal{L}_2$ are the extreme points in the domain of F. This gives

$$\begin{bmatrix} \frac{\partial}{\partial s}J(s_0,h_0) \\ \frac{\partial}{\partial h}J(s_0,h_0) \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial s}F(s_0,h_0) \\ \frac{\partial}{\partial h}F(s_0,h_0) \end{bmatrix} - \gamma_0 \begin{bmatrix} \frac{\partial}{\partial s}G(s_0) \\ \frac{\partial}{\partial h}G(s_0) \end{bmatrix} - \gamma_1 \begin{bmatrix} \frac{\partial}{\partial s}H(h_0) \\ \frac{\partial}{\partial h}H(h_0) \end{bmatrix} = 0, \tag{3.32}$$

which can be rearranged to

$$\begin{bmatrix} \frac{\partial}{\partial s}F(s_0,h_0) \\ \frac{\partial}{\partial h}F(s_0,h_0) \end{bmatrix} = \gamma_0 \begin{bmatrix} \frac{\partial}{\partial s}G(s_0) \\ \frac{\partial}{\partial h}G(s_0) \end{bmatrix} + \gamma_1 \begin{bmatrix} \frac{\partial}{\partial s}H(h_0) \\ \frac{\partial}{\partial h}H(h_0) \end{bmatrix}. \tag{3.33}$$

We have, from Appendix A, that

$$\frac{\partial}{\partial s}G(s_0) = s_0^*, \quad \frac{\partial}{\partial h}G(s_0) = 0, \quad \frac{\partial}{\partial s}H(h_0) = 0, \text{ and } \frac{\partial}{\partial h}H(h_0) = h_0^*. \tag{3.34}$$

Setting $s(t) = x(t) + jy(t)$ and $h(t) = u(t) + jv(t)$, we obtain

$$\frac{\partial}{\partial s}F(s_0, h_0) = \frac{1}{2}\left(\frac{\partial F(s_0, h_0)}{\partial x} - j\frac{\partial F(s_0, h_0)}{\partial y}\right), \tag{3.35}$$

and

$$\frac{\partial}{\partial h}F(s_0, h_0) = \frac{1}{2}\left(\frac{\partial F(s_0, h_0)}{\partial u} - j\frac{\partial F(s_0, h_0)}{\partial v}\right). \tag{3.36}$$

Let $\epsilon \in \mathbb{R}, V \in \mathcal{L}_2$, we obtain the following derivative results.

- $$\begin{aligned}
&F(((x_0 + \epsilon V) + jy_0), h_0) - F(s_0, h_0) \\
=&\iint R_\alpha(\tau, \lambda)((x_0 + \epsilon V) + jy_0)(\tau)h_0(\tau)((x_0 + \epsilon V) - jy_0)(\lambda)h_0^*(\lambda)\, d\tau\, d\lambda - F(s_0, h_0) \\
\approx&\iint R_\alpha(\tau, \lambda)[\epsilon V(\tau)s_0^*(\lambda) + \epsilon V(\lambda)s_0(\tau)]h_0(\tau)h_0^*(\lambda)\, d\tau\, d\lambda \\
=&2\epsilon \iint R_\alpha(\tau, \lambda)h_0(\tau)h_0^*(\lambda)\mathrm{Re}(s_0(\lambda))\, d\lambda\, V(\tau)\, d\tau,
\end{aligned} \tag{3.37}$$

  which implies

$$\frac{\partial F(s_0, h_0)}{\partial x}(\tau) = 2\int R_\alpha(\tau, \lambda)h_0(\tau)h_0^*(\lambda)\mathrm{Re}(s_0(\lambda))\, d\lambda. \tag{3.38}$$

- $$\begin{aligned}
&F(x_0 + j(y_0 + \epsilon V)), h_0) - F(s_0, h_0) \\
=&\iint R_\alpha(\tau, \lambda)(x_0 + j(y_0 + \epsilon V))(\tau)h_0(\tau)(x_0 - j(y_0 + \epsilon V))(\lambda)h_0^*(\lambda)\, d\tau\, d\lambda - F(s_0, h_0) \\
\approx&\iint R_\alpha(\tau, \lambda)[j(\epsilon V(\tau)s_0^*(\lambda) - \epsilon V(\lambda)s_0(\tau))]h_0(\tau)h_0^*(\lambda)\, d\tau\, d\lambda \\
=&2\epsilon \iint R_\alpha(\tau, \lambda)h_0(\tau)h_0^*(\lambda)\mathrm{Im}(s_0(\lambda))\, d\lambda\, V(\tau)\, d\tau,
\end{aligned} \tag{3.39}$$

  which implies

$$\frac{\partial F(s_0, h_0)}{\partial y}(\tau) = 2\int R_\alpha(\tau, \lambda)h_0(\tau)h_0^*(\lambda)\mathrm{Im}(s_0(\lambda))\, d\lambda. \tag{3.40}$$

•

$$
\begin{aligned}
&F(s_0,((u_0+\epsilon V)+jv_0))-F(s_0,h_0)\\
=&\iint R_\alpha(\tau,\lambda)s_0(\tau)((u_0+\epsilon V)+jv_0)(\tau)s_0^*(\lambda)((u_0+\epsilon V)-jv_0)(\lambda)\,d\tau\,d\lambda-F(s_0,h_0)\\
\approx&\iint R_\alpha(\tau,\lambda)[\epsilon V(\tau)h_0^*(\lambda)+\epsilon V(\lambda)h_0(\tau)]s_0(\tau)s_0^*(\lambda)\,d\tau\,d\lambda\\
=&2\epsilon\iint R_\alpha(\tau,\lambda)s_0(\tau)s_0^*(\lambda)\mathrm{Re}(h_0(\lambda))\,d\lambda\,V(\tau)\,d\tau,
\end{aligned}
\tag{3.41}
$$

which implies

$$
\frac{\partial F(s_0,h_0)}{\partial u}(\tau)=2\int R_\alpha(\tau,\lambda)s_0(\tau)s_0^*(\lambda)\mathrm{Re}(h_0(\lambda))\,d\lambda. \tag{3.42}
$$

•

$$
\begin{aligned}
&F(s_0,(u_0+j(v_0+\epsilon V)))-F(s_0,h_0)\\
=&\iint R_\alpha(\tau,\lambda)(u_0+j(v_0+\epsilon V))(\tau)h_0(\tau)(u_0-j(v_0+\epsilon V))(\lambda)h_0^*(\lambda)\,d\tau\,d\lambda-F(s_0,h_0)\\
\approx&\iint R_\alpha(\tau,\lambda)[j(\epsilon V(\tau)h_0^*(\lambda)-\epsilon V(\lambda)h_0(\tau))]s_0(\tau)s_0^*(\lambda)\,d\tau\,d\lambda\\
=&2\epsilon\iint R_\alpha(\tau,\lambda)s_0(\tau)s_0^*(\lambda)\mathrm{Im}(h_0(\lambda))\,d\lambda\,V(\tau)\,d\tau,
\end{aligned}
\tag{3.43}
$$

so

$$
\frac{\partial F(s_0,h_0)}{\partial v}(\tau)=2\int R_\alpha(\tau,\lambda)s_0(\tau)s_0^*(\lambda)\mathrm{Im}(h_0(\lambda))\,d\lambda. \tag{3.44}
$$

Combining the results from (3.34),(3.38), (3.40),(3.42), and (3.44), we obtain

$$
\frac{\partial}{\partial s}F(s_0,h_0)=h_0(\tau)\int R_\alpha(\tau,\lambda)s_0^*(\lambda)h_0^*(\lambda)\,d\lambda=h_0(\tau)L(\tau), \tag{3.45}
$$

and

$$
\frac{\partial}{\partial h}F(s_0,h_0)=s_0(\tau)\int R_\alpha(\tau,\lambda)s_0^*(\lambda)h_0^*(\lambda)\,d\lambda. \tag{3.46}
$$

So, (3.33) implies

$$
\begin{bmatrix}L(\tau)h_0(\tau)\\ L(\tau)s_0(\tau)\end{bmatrix}=\begin{bmatrix}\gamma_0 s_0^*(\tau)\\ \gamma_1 h_0^*(\tau)\end{bmatrix}. \tag{3.47}
$$

From (3.47), we have

$$
L^2(\tau)h_0(\tau)s_0(\tau)=\gamma_0\gamma_1 h_0^*(\tau)s_0^*(\tau), \tag{3.48}
$$

Which further implies

$$|L^2(\tau)||h_0(\tau)||s_0(\tau)| = |\gamma_0\gamma_1||h_0^*(\tau)||s_0^*(\tau)|. \tag{3.49}$$

But

$$|h_0(\tau)| = |h_0^*(\tau)| \quad \text{and} \quad |s_0(\tau)| = |s_0^*(\tau)|, \tag{3.50}$$

so, we have

$$|L(\tau)| = \sqrt{|\gamma_0\gamma_1|} \quad \text{and} \quad |s_0(\tau)| = \left(\sqrt{\left|\frac{\gamma_1}{\gamma_0}\right|}\right)|h_0(\tau)|. \tag{3.51}$$

Because of (3.27), we must have

$$|\gamma_0| = |\gamma_1| \quad \text{so that} \quad |s_0(\tau)| = |h_0(\tau)|. \tag{3.52}$$

That means at the extrema of $F(s,h)$, $s(t)$ and $h(t)$ must be similar in magnitude.

In addition, We have the following three observations:

1. From (3.47), we can observe that $L(\tau)$, $s_0(\tau)$, and $h_0(\tau)$ have the same zeros.

2. From (3.47) and (3.30), we have

$$\begin{aligned} F(s_0,h_0) &= \int L(\lambda)s_0(\lambda)h_0(\lambda)d\lambda \\ &= \int \sqrt{|\gamma_0\gamma_1|}\, s_0(\lambda)s_0^*(\lambda)\, d\lambda = \sqrt{|\gamma_0\gamma_1|}. \end{aligned} \tag{3.53}$$

   That means the SNR of the CP is determined by the square root of the product of the Lagrange multipliers.

3. Let $\{\phi_k\}$ be the set of eigenfunctions of the symmetric kernel R_α with corresponding eigenvalues $\{\kappa_k\}$,

$$\kappa_k\phi_k(\tau) = \int R_\alpha(\tau,\lambda)\phi_k(\lambda)\, d\lambda. \tag{3.54}$$

   Assume that the eigenvalues are distinct for different k. Then from Tricomi [115, Chapter 3],

$$\langle\phi_h,\phi_k\rangle = \int \phi_h(\tau)\phi_k(\tau)d\tau = 0 \quad \text{if} \quad h \neq k. \tag{3.55}$$

   and, $\overline{\text{span}(\{\phi_k\})} = \mathcal{L}_2(a,b)$, with (a,b) being the support of R_α.

Now, let $g(\lambda) = s_0^*(\lambda)h_0^*(\lambda) = \sum a_k \phi_k(\lambda)$, then

$$\begin{aligned} L(\tau) &= \int R_\alpha(\tau, \lambda) g(\lambda) d\lambda \\ &= \int R_\alpha(\tau, \lambda) \sum a_k \phi_k(\lambda) d\lambda = \sum a_k \kappa_k \phi_k(\tau). \end{aligned} \tag{3.56}$$

Because ϕ_k are, orthogonal and κ_k are distinct, we should be able to characterize a_k so that $|L(\tau)| =$ constant. That means we are able to find $f(\lambda)$, and in turn the product $s_0^*(\lambda)h_0^*(\lambda)$. Now, because of (3.52), we may determine $s_0(\lambda)$ and $h_0(\lambda)$ up to complex conjugate pairs.

Chapter 4
Iterative Algorithms for Finding Solutions

The process of finding CPs $s(t)$ and $h(t)$ directly may be prohibitively difficult. In this chapter, we present iterative algorithms for finding these CPs. From the fact that the received signal component is

$$r = \int \alpha(t)s(t)h(t)\ dt, \tag{4.1}$$

which is symmetric in $s(t)$ and $h(t)$, if we are able to characterize $h(t)$ that maximizes the SNR for a given $s(t)$, we are also able to characterize $s(t)$ that maximizes the SNR for a given $h(t)$.

4.1 Iterative Algorithms

Assume that the detection filter is $h_0(t)$, when $s(t)$ is transmitted. Then the received signal at the output of $h_0(t)$ is

$$r = \int [\sum_i \xi_i \alpha_i(t)s(t-\phi_i) + n(t)]h_0(t+\mu)\ dt, \tag{4.2}$$

and the signal and noise components are

$$r_s = \int \sum_i \xi_i \alpha_i(t)s(t-\phi_i)h_0(t+\mu)\ dt, \tag{4.3}$$

and

$$r_n = \int n(t)h_0(t+\mu)\ dt. \tag{4.4}$$

Again, we have $E(|r_n|^2) = N_0/2$. The variance for the signal part is

$$\begin{aligned} E(|r_s|^2) =& E[\int\int (\sum_i \xi_i \alpha_i(\tau)s(\tau-\phi_i))h_0(\tau+\mu)(\sum_j \xi_j^* \alpha_j^*(\lambda)s^*(\lambda-\phi_j))h_0^*(\lambda+\mu)\ d\tau\ d\lambda] \\ =& \sum_i \sum_j \int\int \xi_i \xi_j^* E[\alpha_i(\tau)\alpha_j^*(\lambda)h_0(\tau+\mu)h_0^*(\lambda+\mu)]s(\tau-\phi_i)s^*(\lambda-\phi_j)\ d\tau\ d\lambda. \end{aligned} \tag{4.5}$$

As $\alpha(t)$ and $s_0(t)$ are independent, (4.5) becomes

$$E(|r_s|^2) = \sum_i \sum_j \int\int \xi_i \xi_j^* E[\alpha_i(\tau)\alpha_j^*(\lambda)] E[h_0(\tau+\mu)h_0^*(\lambda+\mu)] s(\tau-\phi_i)s^*(\lambda-\phi_j)\, d\tau\, d\lambda. \tag{4.6}$$

Then, the assumption of independent paths gives

$$E(|r_s|^2) = \sum_i \int\int |\xi_i|^2 E[\alpha_i(\tau)\alpha_i^*(\lambda)] E[h_0(\tau+\mu)h_0^*(\lambda+\mu)] s(\tau-\phi_i)s^*(\lambda-\phi_i)\, d\tau\, d\lambda. \tag{4.7}$$

Defining

$$\begin{aligned} \Upsilon(\tau,\lambda) =& E[h_0(\tau+\mu)h_0^*(\lambda+\mu)] \\ =& \int h_0(\tau+\mu)h_0^*(\lambda+\mu)p(\mu)\, d\mu. \end{aligned} \tag{4.8}$$

Then (4.6) can be rewritten as

$$\begin{aligned} E(|r_s|^2) =& \sum_i \int\int |\xi_i|^2 R_\alpha(\tau,\lambda)\Upsilon(\tau,\lambda)s(\tau-\phi_i)s^*(\lambda-\phi_i)\, d\tau\, d\lambda \\ =& \int\int R_\alpha(\tau-\phi_i,\lambda-\phi_i)\{\sum_i |\xi_i|^2 \Upsilon(\tau+\phi_i,\lambda+\phi_i)\}s(\tau)s^*(\lambda)\, d\tau\, d\lambda. \end{aligned} \tag{4.9}$$

Defining

$$\Theta(\tau,\lambda) = \sum_i |\xi_i|^2 \Upsilon(\tau+\phi_i,\lambda+\phi_i)). \tag{4.10}$$

Again, we have a functional expression of the variance for the signal part.

$$\begin{aligned} E(|r_s|^2) =& \int\int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)s(\tau)s^*(\lambda)\, d\tau\, d\lambda \\ \stackrel{\text{def}}{=}& F(s). \end{aligned} \tag{4.11}$$

We use the method of Lagrange multipliers to derive $s(t)$ that maximizes $F(s)$ under the constraint of $G(s) = \int s^2(t)\, dt = 1$ by setting

$$\frac{d\,F(s_0)}{d\,s} = \gamma \frac{d\,G(s_0)}{d\,s}, \tag{4.12}$$

where $s_0 \in \mathcal{L}_2$ is the extreme point in the domain of F.

For $\epsilon \in \mathbb{R}, V \in \mathcal{L}_2$ with $s(t) = x(t) + jy(t)$, we have

$$\begin{aligned} & F((x_0+\epsilon V)+jy_0) - F(s_0) \\ =& \int\int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)((x_0+\epsilon V)+jy)(\tau)((x_0+\epsilon V)-jy)(\lambda)d\tau d\lambda - F(s_0) \\ \approx& \int\int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)\{\epsilon V(\tau)s_0^*(\lambda) + \epsilon V(\lambda)s_0(\tau)\}\, d\tau\, d\lambda \\ =& 2\epsilon \int\int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)\mathrm{Re}(s_0(\lambda))\, d\lambda\, V(\tau)d\tau \end{aligned} \tag{4.13}$$

both R_α and Θ are symmetric in τ and λ, and

$$\begin{aligned} & F(x_0 + j(y_0 + \epsilon V)) - F(s_0) \\ =& \iint R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)(x_0 + j(y+\epsilon V))(\tau)(x_0 - j(y+\epsilon V))(\lambda)d\tau d\lambda - F(s_0) \\ \approx& \iint R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)\{j[\epsilon V(\tau)s_0^*(\lambda) - \epsilon V(\lambda)s_0(\tau)]\}\, d\tau\, d\lambda \\ =& 2\epsilon \int \int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)\mathrm{Im}(s_0(\lambda))\, d\lambda\, V(\tau)d\tau. \end{aligned} \tag{4.14}$$

That is,

$$\frac{d\,F(s_0)}{d\,s}(\tau) = \int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)s_0^*(\lambda)\, d\lambda. \tag{4.15}$$

Equation (4.15) and (4.12) imply

$$\gamma s_0^*(\tau) = \int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)s_0^*(\lambda)\, d\lambda. \tag{4.16}$$

Additionally,

$$\begin{aligned} E(|r_s|^2) =& F(s_0) = \iint R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)s_0(\tau)s_0^*(\lambda)\, d\tau\, d\lambda \\ =& \int \gamma s_0(\tau)s_0^*(\tau)\, d\tau \\ =& \gamma. \end{aligned} \tag{4.17}$$

Thus to maximize the SNR, we set γ to be the largest eigenvalue of the kernel $R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)$ and set $s_0^*(t)$ to be the corresponding eigenfunction.

By combining the results from (3.22) and (4.17), we obtain an algorithm to find the CP for a multipath fading channel. That is, we iteratively fix the signal function $s(t)$ and the detector function $h(t)$, then evaluate for the corresponding optimal filter function $h_0(t)$ and signal function $s_0(t)$. This is summarized in the following algorithm.

Algorithm 1 *1. Randomly select an initial function $s_0(t)$.*

2. Set $h_0(t)$ so that

$$\gamma_0 h_0^*(\tau) = \int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)h_0^*(\lambda)\, d\lambda, \tag{4.18}$$

where γ_0 is the largest eigenvalue in magnitude of the kernel $R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)$, with $\Psi(\tau,\lambda) = \sum_i |\xi_i|^2 E[s_0(\tau-\phi_i)s_0^*(\lambda-\phi_i)])$. Then $h_0^*(t)$ is the eigenfunction corresponding to γ_0.

3. Set $s_1(t)$ so that for the $h_0(t)$ in the previous step,

$$\gamma_1 s_0^*(\tau) = \int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)s_0^*(\lambda)\, d\lambda, \tag{4.19}$$

where γ_1 is the largest eigenvalue of the kernel $R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)$, with $\Theta(\tau,\lambda) = \sum_i |\xi_i|^2 E[h_0(\tau+\phi_i)h_0^*(\lambda+\phi_i)])$. Then $s_0(t)$ is the corresponding eigenfunction.

4. Repeat step 2 to step 4 a sufficient number of times.

When we try this algorithm with discrete samples of R_α, $s(t)$, and $h(t)$ with the assumption of perfect synchronization using `Matlab`, we find that $|s(t)|^2$ and $|h(t)|^2$ both approach $\delta(t)$. This actually makes sense, because when there is uncertainty in the channel, one would like to transmit information through that channel in as short a time as possible to reduce the uncertainty in the received signal.

Analogous to the characterization of the CP for multiplicative channels which is in the time domain, we may derive the same arguments for frequency selective channels in the frequency domain. This amounts to having a multiplicative factor in the frequency domain. In this case, we will end up with a CP as δ functions in the frequency domain. This agrees with the observation as pointed out by Biglieri [22, p.2636] that "... some recent insights into the peaky nature of capacity-achieving signaling in the realm of broadband time-varying channels...." In addition, a well known capacity-achieving signaling method that attracted much attention is frequency shift keying (FSK), which is "peaky" in the frequency domain.

4.2 Convergence

The precise conditions for which Algorithm 1 converges are unknown. From our experience in implementing this algorithm with discrete approximations, it converges most of the time, at least in the cases when "exact" eigendecompositions are used. On the other hand, proving the monotonic increase of the largest eigenvalue through consecutive iteration of Algorithm 1 for the case of single path fading with no synchronization uncertainty is relatively easy.

Theorem 1 *The sequence of the largest eigenvalue generated through iterations of Algorithm 1 is monotonically increasing for single path fading channels with no synchronization uncertainty.*

Proof:

For single path fading with no synchronization uncertainty, (4.18) becomes

$$\gamma_0 h_0^*(\tau) = \int R_\alpha(\tau,\lambda) s_0(\tau) s_0^*(\lambda) h_0^*(\lambda)\, d\lambda. \tag{4.20}$$

Also, (4.19) becomes

$$\gamma_1 s_0^*(\tau) = \int R_\alpha(\tau,\lambda) h_0(\tau) h_0^*(\lambda) s_0^*(\lambda)\, d\lambda. \tag{4.21}$$

Considering that (4.20) and (4.21) are repeated consecutively by iterative Algorithm 1, we may consolidate these two expressions into one, i.e.,

$$\gamma_k g_k^*(\tau) = \int R_\alpha(\tau,\lambda) g_{k-1}(\tau) g_{k-1}^*(\lambda) g_k^*(\lambda)\, d\lambda, \tag{4.22}$$

where γ_k is the largest eigenvalue for the kth iteration, and g_k replaces s and h alternatively. Now, we have from (4.22) and (3.27) that

$$\begin{aligned}
\gamma_k &= \int g_k(\tau) g_k^*(\tau)\, d\tau \\
&= \int\int R_\alpha(\tau,\lambda) g_{k-1}(\tau) g_{k-1}^*(\lambda) g_k(\tau) g_k^*(\lambda)\, d\lambda\, d\tau \\
&\geq \int\int R_\alpha(\tau,\lambda) g_{k-1}(\tau) g_{k-1}^*(\lambda) g_{k-2}(\tau) g_{k-2}^*(\lambda)\, d\lambda\, d\tau \\
&= \gamma_{k-1}.
\end{aligned} \tag{4.23}$$

This completes the proof.

We are interested in finding CPs that maximize the SNR, and the SNR is related to the largest eigenvalue. So, the variation of the largest eigenvalue through the iteration can serve as an indicator of the progress of the algorithm.

4.3 Introduction of Frequency Constraints

From section 4.1, we found that CPs of single path fading channels with no synchronization uncertainty are δ distributions. However, δ distributions are not practical signaling waveforms. One possible approach for designing practical CPs that are "optimal" yet producible is to introduce frequency restrictions to $s(t)$ and/or $h(t)$. In addition to requiring $s(t)$ and $h(t)$ to be unit powered, we impose restrictions on the frequency spread of $s(t)$ and/or $h(t)$. Let $S(f) = \mathcal{F}(s(t))$ and $H(f) = \mathcal{F}(h(t))$ denote the Fourier transforms. We introduce the functional $W_s(s)$ and $W_h(h)$ as measures of the energy of $s(t)$ and $h(t)$, respectively in the frequency range $(-f, f)Hz$. Specifically, we let

$$\begin{aligned} W_s(s) &= \int_{-f}^{f} S(\eta)S^*(\eta)\, d\eta \\ &= \int_{-\infty}^{\infty} [\Pi(\frac{\eta}{2f})S(\eta)]S^*(\eta)\, d\eta \\ &= 2f \int_{-\infty}^{\infty} [s(t) * \mathrm{sinc}(2\pi f t)]s^*(t)\, dt \\ &= 2f \int_{-\infty}^{\infty} s^*(t) \int_{-\infty}^{\infty} s(\tau)\,\mathrm{sinc}(2\pi f(t-\tau))\, d\tau\, dt \\ &= 2f \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{sinc}(2\pi f(\tau-\lambda))s(\tau)s^*(\lambda)\, d\tau\, d\lambda, \end{aligned} \tag{4.24}$$

and similarly

$$\begin{aligned} W_h(h) &= \int_{-f}^{f} H(f)H^*(f)\, df \\ &= 2f \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{sinc}(2\pi f(\tau-\lambda))h(\tau)h^*(\lambda)\, d\tau\, d\lambda. \end{aligned} \tag{4.25}$$

We reconstruct the Lagrangian in (3.14) and (4.12)

$$(1-\beta)\begin{bmatrix} \frac{\partial}{\partial s}F(s_0) \\ \frac{\partial}{\partial h}F(h_0) \end{bmatrix} + \beta \begin{bmatrix} \frac{\partial}{\partial s}W_s(s_0) + \frac{\partial}{\partial s}W_h(h_0) \\ \frac{\partial}{\partial h}W_s(s_0) + \frac{\partial}{\partial h}W_h(h_0) \end{bmatrix} = \gamma_0 \begin{bmatrix} \frac{\partial}{\partial s}G(s_0) \\ \frac{\partial}{\partial h}G(s_0) \end{bmatrix} + \gamma_1 \begin{bmatrix} \frac{\partial}{\partial s}H(h_0) \\ \frac{\partial}{\partial h}H(h_0) \end{bmatrix}. \tag{4.26}$$

where $\beta \in [0,1]$ is a weighing factor between the SNR and frequency containment. As $W_s(s)$ is independent of h, we have $\frac{\partial}{\partial h}W_s(s_0) = 0$. Again, let $s(t) = x(t) + jy(t)$, then

evaluate

$$\begin{aligned}
&W_s((x_0+\epsilon V)+jy_0)-W_s(s_0)\\
=&2f\int\int \mathrm{sinc}(2\pi f(\tau-\lambda))((x_0+\epsilon V)+jy_0)(\tau)((x_0+\epsilon V)-jy)(\lambda)\,d\lambda\,d\tau - W_s(s_0)\\
\approx&2f\epsilon\int\int \mathrm{sinc}(2\pi f(\tau-\lambda))[s_0(\tau)V(\lambda)+s_0^*(\lambda)V(\tau)]\,d\lambda\,d\tau\\
=&4f\epsilon\int\int \mathrm{sinc}(2\pi f(\tau-\lambda))\mathrm{Re}(s_0(\lambda))V(\tau)\,d\lambda\,d\tau,
\end{aligned} \tag{4.27}$$

and

$$\begin{aligned}
&W_s(x_0+j(y_0+\epsilon V))-W_s(s_0)\\
=&2f\int\int \mathrm{sinc}(2\pi f(\tau-\lambda))(x_0+j(y_0+\epsilon V))(\tau)(x_0-j(y+\epsilon V))(\lambda)\,d\lambda\,d\tau - W_s(s_0)\\
\approx&2f\epsilon\int\int \mathrm{sinc}(2\pi f(\tau-\lambda))[j(s_0^*(\lambda)V(\tau)-s_0(\tau)V(\lambda))]\,d\lambda\,d\tau\\
=&4f\epsilon\int\int \mathrm{sinc}(2\pi f(\tau-\lambda))\mathrm{Im}(s_0(\lambda))V(\tau)\,d\lambda\,d\tau.
\end{aligned} \tag{4.28}$$

We have

$$\frac{\partial}{\partial s}W_s(s_0)=2f\int \mathrm{sinc}(2\pi f(\tau-\lambda))s_0^*(\lambda)\,d\lambda=K(\tau). \tag{4.29}$$

Similarly, we have $\dfrac{\partial}{\partial s}W_h(h_0)=0$, and

$$\frac{\partial}{\partial h}W_h(h_0)=2f\int \mathrm{sinc}(2\pi f(\tau-\lambda))h_0^*(\lambda)\,d\lambda=M(\tau). \tag{4.30}$$

As a result, (4.26) becomes

$$\begin{bmatrix}(1-\beta)\int R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)s_0^*(\lambda)\,d\lambda+\beta K(\tau)\\ (1-\beta)\int R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)h_0^*(\lambda)\,d\lambda+\beta M(\tau)\end{bmatrix}=\begin{bmatrix}\gamma_0 s_0^*(\tau)\\ \gamma_1 h_0^*(\tau)\end{bmatrix}. \tag{4.31}$$

If we use the iterative approach, we have the following algorithm:

Algorithm 2 *1. Randomly select an initial function $s_0(t)$.*

2. Set $h_0(t)$ so that

$$\gamma_1 h_0^*(\tau)=\int[(1-\beta)R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)+2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]h_0^*(\lambda)\,d\lambda, \tag{4.32}$$

γ_1 is the largest eigenvalue of the kernel $[(1-\beta)R_\alpha(\tau,\lambda)\Psi(\tau,\lambda)+2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$, while $h_0^*(t)$ is the corresponding eigenfunction.

3. Set $s_0(t)$ so that for $h_0(t)$ from the last step,

$$\gamma_0 s_0^*(\tau) = \int [(1-\beta)R_\alpha(\tau,\lambda)\Theta(\tau,\lambda) + 2\beta f \operatorname{sinc}(2\pi f(\tau-\lambda))] s_0^*(\lambda)\, d\lambda, \qquad (4.33)$$

γ_0 is the largest eigenvalue in of the kernel $[(1-\beta)R_\alpha(\tau,\lambda)\Theta(\tau,\lambda)+2\beta f \operatorname{sinc}(2\pi f(\tau-\lambda))]$ while $s_0^*(t)$ is the corresponding eigenfunction.

4. Repeat steps 2, 3 and 4 a sufficient number of times.

Now, Algorithm 1 is a special case of Algorithm 2.

4.4 Signal Sets

To communicate over a fast fading channel whose amplitude may change sign instantly, amplitude and phase modulation are inappropriate, and something akin to orthogonal signaling is necessary. We must therefore find at least two sets of signal/filter pairs $(s_0(t), h_0(t))$ and $(s_1(t), h_1(t))$, such that under the channel influence $\alpha(t)$, $s_i(t)$ and $h_j(t)$ have a minimal correlation for $i \neq j$. This minimal correlation criterion may be fulfilled by a direction application of the CPs generated with Algorithm 2 in pulse-position modulation (PPM). Another more general approach requires an extension to Algorithm 2.

From Algorithm 2, we know that the minimum correlation between $h_j(t)$ and $s_i(t)$ can be obtained by setting $h_j(t)$ to be the eigenfunction corresponding to the smallest eigenvalue while we generate $h_i(t)$ for given $s_i(t)$, at the same time setting $s_j(t)$ to be the eigenfunction corresponding to the smallest eigenvalue while we generate $s_i(t)$ for given $h_i(t)$. However, the kernel must be changed to enforce frequency containment. So, while finding the eigenfunctions corresponding to the smallest eigenvalues, we use the kernel of the inverse operator ($\mathcal{I}[\cdot]$, Appendix B) of (4.29) and (4.30) in place of $2f \operatorname{sinc}(2\pi f(\tau-\lambda))$ in (4.31) for frequency containment.

We suggest the following procedure: Given a transmitted signal $s_0(t)$, let the corresponding filter $\hat{h}_0(t)$ be the eigenfunction having largest eigenvalue of the kernel $(1-\beta)R_\alpha(\tau,\lambda)\Psi(\tau,\lambda) + 2\beta f \operatorname{sinc}(2\pi f(\tau-\lambda))$ as defined before. At the same time, set $\tilde{h}_1(t)$ to be the eigenfunction with smallest eigenvalue of the kernel $(1-\beta)R_\alpha(\tau,\lambda)\Psi(\tau,\lambda) +$

$\beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$. Similarly, given a detection filter $h_0(t)$, we let the corresponding signal $\hat{s}_0(t)$ be the eigenfunction having largest eigenvalue of kernel $(1-\beta)R_\alpha(\tau,\lambda)\Phi(\tau,\lambda)+2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))$. At the same time, set $\tilde{s}_1(t)$ to be the eigenfunction with smallest eigenvalue of the kernel $(1-\beta)R_\alpha(\tau,\lambda)\Phi(\tau,\lambda)+\beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$. Use the same method to obtain $\hat{s}_1(t)$, $\hat{h}_1(t)$, $\tilde{s}_0(t)$, and $\tilde{h}_0(t)$. This procedure will be carried out repeatedly: for each new iteration, set $s_0(t)=\hat{s}_0(t)+\tilde{s}_0(t)$, $s_1(t)=\hat{s}_1(t)+\tilde{s}_1(t)$, $h_0(t)=\hat{h}_0(t)+\tilde{h}_0(t)$, and $h_1(t)=\hat{h}_1(t)+\tilde{h}_1(t)$, then normalize these functions to unit energy. The procedure is summarized in Fig. 4.1.

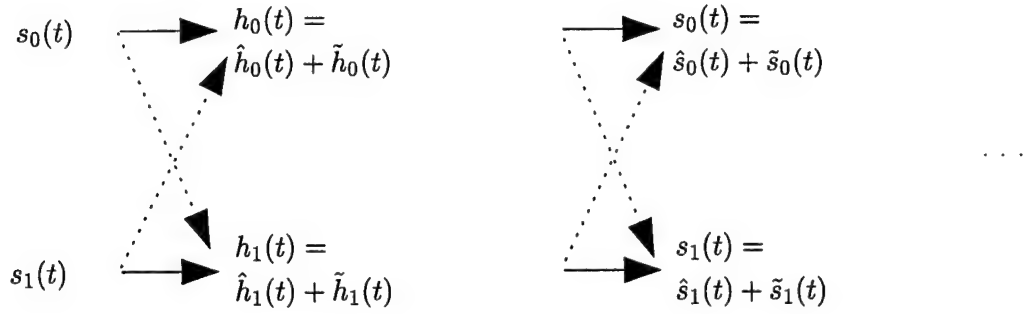


Fig. 4.1: Procedure for generating signal/filter pairs.

In the above procedure, $s_0(t)$ and $h_0(t)$ form a CP, while $s_1(t)$ and $h_1(t)$ form another CP. Further, the correlation between $s_0(t)$ and $h_1(t)$, and that between $s_1(t)$ and $h_0(t)$ are small. We call CPs with minimal correlation as **conjugate-consonant pairs** (CCPs).

Due to the averaging in the generation process of the signal/filter pairs, the CCPs vary a lot, both with the waveforms and the SNR in terms of the largest eigenvalue. However, during the process, we come across many "good" pairs which have high SNR.

Algorithm 3 *1. Randomly select initial functions $s_0(t)$ and $s_1(t)$.*

2. Set $h_{00}(t)$ so that

$$\gamma_{00}h_{00}^*(\tau)=\int[(1-\beta)R_\alpha(\tau,\lambda)\Psi_0(\tau,\lambda)+2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]h_{00}^*(\lambda)\,d\lambda, \quad (4.34)$$

γ_{00} is the largest eigenvalue of the kernel $[(1-\beta)R_\alpha(\tau,\lambda)\Psi_0(\tau,\lambda)+2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$, while $h_{00}^*(t)$ is the corresponding eigenfunction.

3. *Set* h_{01} *so that*

$$\xi_{01}h_{01}^*(\tau) = \int [(1-\beta)R_\alpha(\tau,\lambda)\Psi_0(\tau,\lambda) + \beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]]h_{01}^*(\lambda)\,d\lambda, \quad (4.35)$$

ξ_{11} *is the smallest eigenvalue of the kernel* $[(1-\beta)R_\alpha(\tau,\lambda)\Psi_0(\tau,\lambda)+\beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]]$, *while* $h_{01}^*(t)$ *is the corresponding eigenfunction.*

4. *Set* $h_{11}(t)$ *so that*

$$\gamma_{11}h_{11}^*(\tau) = \int [(1-\beta)R_\alpha(\tau,\lambda)\Psi_1(\tau,\lambda) + 2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]h_{10}^*(\lambda)\,d\lambda, \quad (4.36)$$

γ_{11} *is the largest eigenvalue of the kernel* $[(1-\beta)R_\alpha(\tau,\lambda)\Psi_1(\tau,\lambda)+2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$, *while* $h_{11}^*(t)$ *is the corresponding eigenfunction.*

5. h_{10} *so that*

$$\xi_{10}h_{10}^*(\tau) = \int [(1-\beta)R_\alpha(\tau,\lambda)\Psi_1(\tau,\lambda) + \beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]]h_{10}^*(\lambda)\,d\lambda, \quad (4.37)$$

ξ_{10} *is the smallest eigenvalue of the kernel* $[(1-\beta)R_\alpha(\tau,\lambda)\Psi_1(\tau,\lambda)+\beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]]$, *while* $h_{10}^*(t)$ *is the corresponding eigenfunction.*

6. *Find the average, i.e.,*

$$h_0(t) = h_{00}(t) + h_{10}(t), \quad (4.38)$$

and

$$h_1(t) = h_{01}(t) + h_{11}(t). \quad (4.39)$$

7. *Normalize* $h_0(t)$ *and* $h_1(t)$.

8. *Set* $s_{00}(t)$ *so that*

$$\gamma_{00}s_{00}^*(\tau) = \int [(1-\beta)R_\alpha(\tau,\lambda)\Theta_0(\tau,\lambda) + 2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]s_{00}^*(\lambda)\,d\lambda, \quad (4.40)$$

γ_{00} *is the largest eigenvalue of the kernel* $[(1-\beta)R_\alpha(\tau,\lambda)\Theta_0(\tau,\lambda)+2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$, *while* $s_{00}^*(t)$ *is the corresponding eigenfunction.*

9. *Set* s_{01} *so that*

$$\xi_{01} s_{01}^*(\tau) = \int [(1-\beta) R_\alpha(\tau,\lambda)\Theta_0(\tau,\lambda) + \beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]] s_{01}^*(\lambda)\, d\lambda, \quad (4.41)$$

ξ_{11} *is the smallest eigenvalue of the kernel* $[(1-\beta)R_\alpha(\tau,\lambda)\Theta_0(\tau,\lambda)+\beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]]$, *while* $s_{01}^*(t)$ *is the corresponding eigenfunction.*

10. *Set* $s_{11}(t)$ *so that*

$$\gamma_{11} s_{11}^*(\tau) = \int [(1-\beta) R_\alpha(\tau,\lambda)\Theta_1(\tau,\lambda) + 2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))] s_{10}^*(\lambda)\, d\lambda, \quad (4.42)$$

γ_{11} *is the largest eigenvalue of the kernel* $[(1-\beta)R_\alpha(\tau,\lambda)\Theta_1(\tau,\lambda)+2\beta f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$, *while* $s_{11}^*(t)$ *is the corresponding eigenfunction.*

11. s_{10} *so that*

$$\xi_{10} s_{10}^*(\tau) = \int [(1-\beta) R_\alpha(\tau,\lambda)\Theta_1(\tau,\lambda) + \beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]] s_{10}^*(\lambda)\, d\lambda, \quad (4.43)$$

ξ_{10} *is the smallest eigenvalue of the kernel* $[(1-\beta)R_\alpha(\tau,\lambda)\Theta_1(\tau,\lambda)+\beta\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]]$, *while* $s_{10}^*(t)$ *is the corresponding eigenfunction.*

12. *Find the average, i.e.,*

$$s_0(t) = s_{00}(t) + s_{10}(t), \quad (4.44)$$

and

$$s_1(t) = s_{01}(t) + s_{11}(t). \quad (4.45)$$

13. *Normalize* $s_0(t)$ *and* $s_1(t)$.

14. *Repeat steps 2 to 13 a sufficient number of times. Save those CCPs with high SNR along the iteration.*

Chapter 5
The Study of Probability of Error

5.1 Introduction

In conjunction with the algorithms suggested in the previous chapter, we would like to develop a method to evaluate the performance of the derived CPs in terms of SNR versus probability of error, $P(e)$. So, we may compare the performance between the CPs generated through those procedures in chapter 4 and other traditional waveforms.

Up to this point, we have not imposed any specification for the fading process $\alpha(t)$, so the principles we derived can be applied to general fading channels. However, for the following derivation, we will assume the fading process is a stationary band-limited Gaussian process to facilitate our development. The result we obtain will be applicable to Rayleigh fading channels, which are the most common fading channel models.

5.2 Quadratic Forms of Complex Normal Vectors

Let $\{v_0, v_1, \cdots, v_{N-1}\}$ be complex random variables with real and imaginary parts, $v_k = x_k + jy_k$, which are normally distributed and satisfy the following relationships:

$$E[(x_i - \bar{x}_i)(x_j - \bar{x}_j)] = E[(y_i - \bar{y}_i)(y_j - \bar{y}_j)] \tag{5.1}$$

and

$$E[(x_i - \bar{x}_i)(y_j - \bar{y}_j)] = -E[(x_j - \bar{x}_j)(y_i - \bar{y}_i)], \tag{5.2}$$

where $\bar{x}_k = E(x_k)$ and $\bar{y}_k = E(y_k)$. Define the complex normal vector $\mathbf{v}$ as

$$\mathbf{v} = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{bmatrix}, \tag{5.3}$$

with mean $E(\mathbf{v}) = 0$, and covariance $M = E[\mathbf{v}\mathbf{v}^H]$. A quadratic form of $\mathbf{v}$ is defined as

$$g = \mathbf{v}^H Q \mathbf{v}, \tag{5.4}$$

where Q is any Hermitian symmetric matrix.

According to Turin [116], the probability density function of g, $p(g)$, has a characteristic function

$$\Phi(t) = \prod_{n=1}^{N} \frac{1}{1 - jt\varrho_n}, \tag{5.5}$$

where ϱ_k are the eigenvalues of MQ. To simplify our development, we assume that all ϱ_k are distinct.[1] Since $\varrho_k = 0$ does not affect $\Phi(t)$, we assume that all the eigenvalues are nonzero.

The probability density function of g is

$$p(g) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(t) \exp(-jtg)\, dt, \tag{5.7}$$

which can be evaluated by using the residue theorem,

$$\begin{aligned} p(g) =& \frac{1}{2\pi} \int_C \Phi(z) \exp(-jzg)\, dz \\ =& j \sum \{\text{residues enclosed in the counterclockwise contour}\}, \end{aligned} \tag{5.8}$$

where C is a contour in the complex plane containing the real axis. The integrand in (5.8) can be rearranged as

$$\begin{aligned} \Phi(z) \exp(-jzg) &= \left\{ \prod_{n=1}^{N} \frac{1}{1 - jz\varrho_n} \right\} \exp(-jzg) \\ &= \left\{ \prod_{n=1}^{N} \frac{j\varrho_n^{-1}}{z + j\varrho_n^{-1}} \right\} \exp(-jzg). \end{aligned} \tag{5.9}$$

For each isolated singular point $-j\varrho_i^{-1}$, we can represent the integrand by Laurent series

$$\begin{aligned} &\Phi(z) \exp(-jzg) \\ &= \left\{ \prod_{n=1, n\neq i}^{N} \frac{j\varrho_n^{-1}}{z + j\varrho_n^{-1}} \right\} \frac{j\varrho_i^{-1}}{z + j\varrho_i^{-1}} \exp(-jzg) \\ &= \left\{ \prod_{n=1, n\neq i}^{N} \frac{j\varrho_n^{-1}}{z + j\varrho_n^{-1}} \right\} \left\{ j\varrho_i^{-1} \exp(-\varrho_i^{-1} g) \left[\frac{1}{z + j\varrho_i^{-1}} (1 + (z + j\varrho_i^{-1}) + \frac{(z + j\varrho_i^{-1})^2}{2} + \cdots) \right] \right\} \end{aligned} \tag{5.10}$$

---

[1] When there are repeated eigenvalues, the characteristic function is

$$\Phi(t) = \prod_{n} \frac{1}{(1 - jt\varrho_n)^{\mu_n}}, \tag{5.6}$$

where μ_n is the number of replications of ϱ_n. There will be some minor modifications to our derivation when $\mu_n \neq 1$ for some n. To simplify our discussion, we assume there are no duplicated eigenvalues.

Hence, the residues of $\Phi(z)\exp(-jgz)$ are

$$\text{Res}\{\Phi(z)\exp(-jgz), -j\varrho_i^{-1}\} = j\varrho_i^{-1}\left\{\prod_{n=1,n\neq i}^{N}\frac{\varrho_i}{\varrho_i-\varrho_n}\right\}\exp(-\varrho_i^{-1}g) \tag{5.11}$$

A contour map is shown in Fig. 5.1. In this figure, there are two possible closed loops, I_1+I_2 and I_1+I_3. Which contour we use depends on the sign of g: e^{-jzg} must be bounded. For $z=x+jy$, we have the following cases:

1. $g\geq 0$: $e^{-jgz}=e^{-jgx}\cdot e^{gy}$ implies that we need to take $y\leq 0$ and the loop $C=I_1+I_3$. By the residue theorem,

$$p(g)=\sum_{\varrho_i>0}\varrho_i^{-1}\left\{\prod_{n=1,n\neq i}^{N}\frac{\varrho_i}{\varrho_i-\varrho_n}\right\}\exp(-\varrho_i^{-1}g) \tag{5.12}$$

2. $g<0$: $e^{-jgz}=e^{-jgx}\cdot e^{gy}$ implies that we need to take $y\geq 0$ and $C=I_1+I_2$. So,

$$p(g)=-\sum_{\varrho_i<0}\varrho_i^{-1}\left\{\prod_{n=1,n\neq i}^{N}\frac{\varrho_i}{\varrho_i-\varrho_n}\right\}\exp(-\varrho_i^{-1}g) \tag{5.13}$$

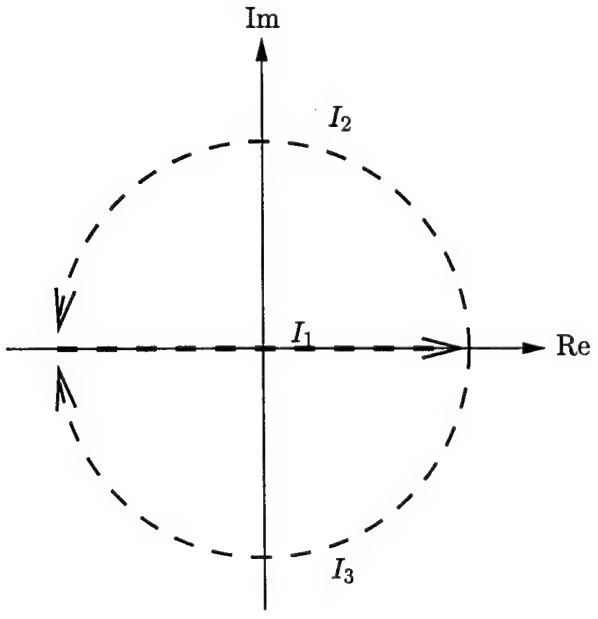


Fig. 5.1: Contour map.

5.3 Cumulative Distribution Function

Theorem 2 *Let* $\mathbf{v}$ *be a complex normal vector with mean* $E(\mathbf{v}) = 0$, *and covariance* $M = E[\mathbf{v}\mathbf{v}^H]$. *Further let* Q *be any Hermitian matrix. Then the quadratic form* $g = \mathbf{v}^H Q \mathbf{v}$ *of* $\mathbf{v}$ *has a cumulative distribution function of*

$$F_g(\xi) = \begin{cases} \sum_{\varrho_i<0} \left\{ \prod_{n=1,n\neq i}^{N} \frac{\varrho_i}{\varrho_i-\varrho_n} \right\} \exp(-\xi\varrho_i^{-1}) & \text{if} \quad \xi < 0 \\ 1 - \sum_{\varrho_i>0} \left\{ \prod_{n=1,n\neq i}^{N} \frac{\varrho_i}{\varrho_i-\varrho_n} \right\} \exp(-\xi\varrho_i^{-1}) & \text{if} \quad \xi \geq 0, \end{cases} \tag{5.14}$$

where ϱ_k *are the eigenvalues of* MQ.

Proof:

The computation of $F_g(\xi)$, the cumulative distribution function of g, falls into the following two cases, following (5.12) and (5.13).

1. $\xi < 0$:

$$\begin{aligned} F_g(\xi) &= \int_{-\infty}^{\xi} p(g)\, dg \\ &= -\int_{-\infty}^{\xi} \sum_{\varrho_i<0} \varrho_i^{-1} \left\{ \prod_{n=1,n\neq i}^{N} \frac{\varrho_i}{\varrho_i-\varrho_n} \right\} \exp(-\varrho_i^{-1} g)\, dg \\ &= -\sum_{\varrho_i<0} \varrho_i^{-1} \left\{ \prod_{n=1,n\neq i}^{N} \frac{\varrho_i}{\varrho_i-\varrho_n} \right\} \int_{-\infty}^{\xi} \exp(-\varrho_i^{-1} g)\, dg \\ &= \sum_{\varrho_i<0} \left\{ \prod_{n=1,n\neq i}^{N} \frac{\varrho_i}{\varrho_i-\varrho_n} \right\} \exp(-\xi\varrho_i^{-1}). \end{aligned} \tag{5.15}$$

2. $\xi \geq 0$:

$$\begin{aligned} F_g(\xi) &= \int_{-\infty}^{\xi} p(g)\, dg \\ &= 1 - \int_{\xi}^{\infty} p(g)\, dg \\ &= 1 - \int_{\xi}^{\infty} \sum_{\varrho_i>0} \varrho_i^{-1} \left\{ \prod_{n=1,n\neq i}^{N} \frac{\varrho_i}{\varrho_i-\varrho_n} \right\} \exp(-\varrho_i^{-1} g)\, dg \\ &= 1 - \sum_{\varrho_i>0} \varrho_i^{-1} \left\{ \prod_{n=1,n\neq i}^{N} \frac{\varrho_i}{\varrho_i-\varrho_n} \right\} \int_{\xi}^{\infty} \exp(-\varrho_i^{-1} g)\, dg \\ &= 1 - \sum_{\varrho_i>0} \left\{ \prod_{n=1,n\neq i}^{N} \frac{\varrho_i}{\varrho_i-\varrho_n} \right\} \exp(-\xi\varrho_i^{-1}). \end{aligned} \tag{5.16}$$

5.4 A Partition of Unity

The following results do not relate directly to our development of CPs. We present them here as they are very interesting and also very useful in the field of digital communication. When $\xi = 0$, we have the following corollary immediately from (5.15) and (5.16).

Corollary 1 *Let* $\{\varrho_1, \varrho_2, \cdots, \varrho_N\}$ *be distinct real numbers, then*

$$\sum_{i=1}^{N} \prod_{n=1, n\neq i}^{N} \frac{\varrho_i}{\varrho_i - \varrho_n} = 1. \tag{5.17}$$

The above derivation of the cumulative distribution function proves that this corollary is true for real sequences of $\{\varrho_1, \varrho_2, \cdots, \varrho_N\}$. This result can be extended to include distinct complex sequences [117].

The above partition of unity can also be expressed by the following beautiful forms:

1. Let $\varrho_i = i\varrho_1$, then

$$\begin{aligned} &\sum_{i=1}^{N} \prod_{n=1, n\neq i}^{N} \frac{\varrho_i}{\varrho_i - \varrho_n} = \sum_{i=1}^{N} \prod_{n=1, n\neq i}^{N} \frac{i}{i-n} = 1 \\ \Rightarrow &\frac{1}{(N-1)!} \sum_{i=1}^{N} (-1)^{N-1} \binom{N-1}{i-1} i^{N-1} = 1. \end{aligned} \tag{5.18}$$

2. When $\varrho_i = \frac{1}{i}\varrho_1$, then

$$\begin{aligned} &\sum_{i=1}^{N} \prod_{n=1, n\neq i}^{N} \frac{\varrho_i}{\varrho_i - \varrho_n} = \sum_{i=1}^{N} \prod_{n=1, n\neq i}^{N} \frac{n}{n-i} = 1 \\ \Rightarrow &N \sum_{i=1}^{N} (-1)^{N-i+1} \binom{N-1}{i-1} \frac{1}{i} = 1. \end{aligned} \tag{5.19}$$

3. For $\varrho_i = x^i$ where $x \neq \{0, 1\}$,

$$\sum_{i=1}^{N} \prod_{n=1, n\neq i}^{N} \frac{1}{1 - x^{n-i}} = 1, \tag{5.20}$$

4. For $\varrho_i = i^x$, $x \neq 0$,

$$\sum_{i=1}^{N} \prod_{n=1, n\neq i}^{N} \frac{1}{1-(n/i)^x} = 1. \tag{5.21}$$

Other than being pretty in expression, the above partition of unity can also be used to simplify many expressions presented in the literature, for example [21, p.802(14-5-27)] and [92, p.274(9.26)].

5.5 Quadratic Receiver

A quadratic receiver is a receiver that employs a filter bank to obtain a received signal-plus-noise vector, then determines which signal is transmitted out of a finite number of possible signals by considering the signal-plus-noise vector as a whole. Assume the channel is $C(\tau, t)$ with AWGN and may be either a fading or a nonfading channel. There are K possible transmitted signals $\{s_0(t), s_1, \cdots, s_{K-1}(t)\}$, and there are N filters in the receiver filter bank $\{h_0(t), h_1(t), \cdots, h_{N-1}(t)\}$. Fig. 5.2 shows a quadratic receiver. Let $r_k(t)$ be the received signal before the receiver when $s_k(t)$ is transmitted through channel $C(\tau, t)$ with AWGN $n(t)$. Then

$$r_k(t) = c(\tau, t) * s_k(t) + n(t), \tag{5.22}$$

where $n(t)$ has a power spectral density of $N_0/2$. Define

$$\begin{aligned} r_{k_i} &= \int r_k(t) h_i(t)\, dt \\ &= \int c(\tau, t) * s_k(t) h_i(t)\, dt + \int n(t) h_i(t)\, dt \\ &= \int \sum_j \xi_j \alpha_j(t) s_k(t - \phi_j) h_i(t)\, dt + \int n(t) h_i(t)\, dt, \end{aligned} \tag{5.23}$$

The signal-plus-noise vector when $s_k(t)$ is sent is

$$\mathbf{r}_k = \begin{bmatrix} r_{k_0} \\ r_{k_1} \\ \vdots \\ r_{k_{N-1}} \end{bmatrix}. \tag{5.24}$$

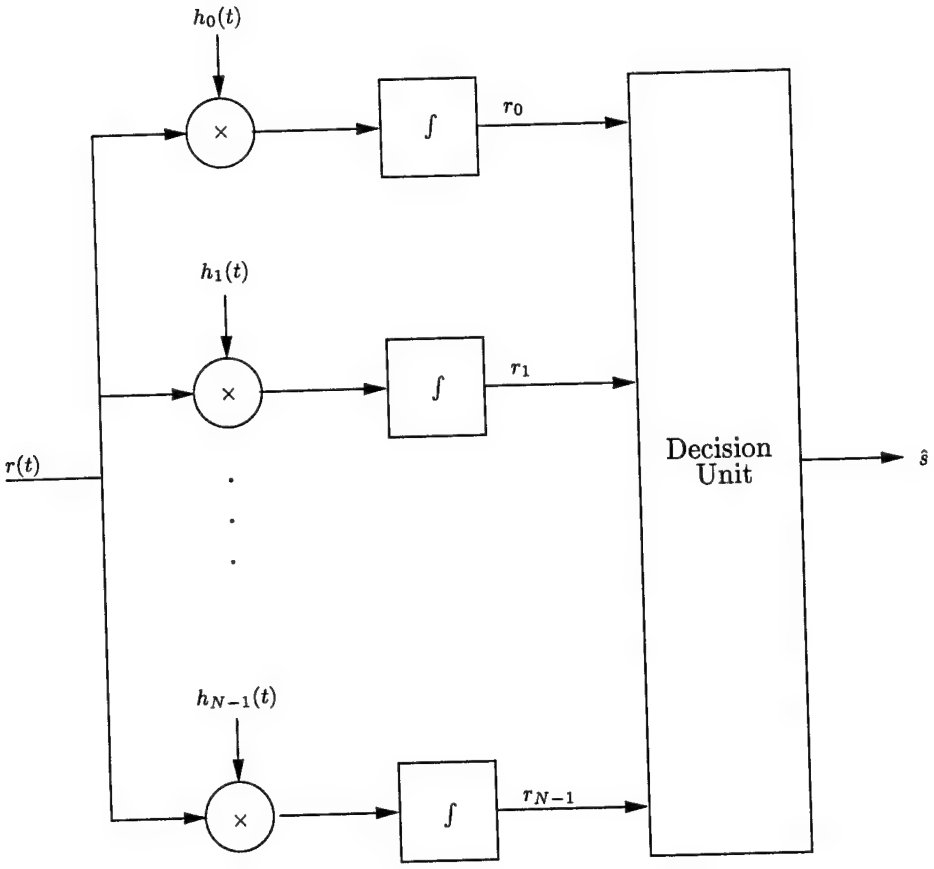


Fig. 5.2: Detection unit of quadratic receiver.

Let $\bar{\mathbf{r}}_k = E(\mathbf{r}_k)$. The covariance matrix of the sample vector when $s_k(t)$ is transmitted is defined as

$$\begin{aligned} R_k =& E[(\mathbf{r}_k - \bar{\mathbf{r}}_k)(\mathbf{r}_k - \bar{\mathbf{r}}_k)^H] \\ =& E(\mathbf{r}_k\mathbf{r}_k^H - \mathbf{r}_k\bar{\mathbf{r}}_k^H - \bar{\mathbf{r}}_k\mathbf{r}_k^H + \bar{\mathbf{r}}_k\bar{\mathbf{r}}_k^H) \\ =& E(\mathbf{r}_k\mathbf{r}_k^H) - E(\mathbf{r}_k)\bar{\mathbf{r}}_k^H - \bar{\mathbf{r}}_k E(\mathbf{r}_k^H) + \bar{\mathbf{r}}_k\bar{\mathbf{r}}_k^H \\ =& E(\mathbf{r}_k\mathbf{r}_k^H) - \bar{\mathbf{r}}_k\bar{\mathbf{r}}_k^H. \end{aligned} \tag{5.25}$$

We have

$$\bar{\mathbf{r}}_k = \begin{bmatrix} \bar{r}_{k_0} \\ \bar{r}_{k_1} \\ \vdots \\ \bar{r}_{k_{N-1}} \end{bmatrix}, \tag{5.26}$$

where

$$\bar{r}_{k_i} = \int E[C(\tau,t)] * s_k(t)h_i(t)\ dt + \int E[n(t)h_i(t)]\ dt, \tag{5.27}$$

which will be zero for fast fading.

To simplify our discussion, we assume that $K = 2$, and let

$$C_k = \begin{bmatrix} \bar{r}_{k_0}\bar{r}_{k_0}^* & \bar{r}_{k_0}\bar{r}_{k_1}^* \\ \bar{r}_{k_1}\bar{r}_{k_0}^* & \bar{r}_{k_1}\bar{r}_{k_1}^* \end{bmatrix} \tag{5.28}$$

which is zero for fast fading. Also,

$$\begin{aligned} E(\mathbf{r}_k\mathbf{r}_k^H) =& E(\begin{bmatrix} r_{k_0}r_{k_0}^* & r_{k_0}r_{k_1}^* \\ r_{k_1}r_{k_0}^* & r_{k_1}r_{k_1}^* \end{bmatrix}) \\ =& \begin{bmatrix} r_{k00} & r_{k01} \\ r_{k10} & r_{k11} \end{bmatrix} + \frac{N_0}{2}\begin{bmatrix} 1 & \int h_0(t)h_1(t)\ dt \\ \int h_1(t)h_0(t)\ dt & 1 \end{bmatrix}, \end{aligned} \tag{5.29}$$

where

$$\begin{aligned}
r_{k00} =& E[\int (\sum_i \xi_i \alpha_i(\tau) s_k(\tau - \phi_i - \mu)) h_0(\tau)\, d\tau \int (\sum_j \xi_j^* \alpha_i^*(\lambda) s_k^*(\lambda - \phi_j + \mu)) h_0^*(\lambda)\, d\lambda] \\
=& E[\int \int (\sum_i \xi_i \alpha_i(\tau) s_k(\tau - \phi_i - \mu))(\sum_j \xi_j^* \alpha_j^*(\lambda) s_k^*(\lambda - \phi_j - \mu)) h_0(\tau) h_0^*(\lambda)\, d\tau\, d\lambda] \\
=& \sum_i \sum_j E[\int \int (\xi_i \alpha_i(\tau) s_k(\tau - \phi_i - \mu))(\xi_j^* \alpha_j^*(\lambda) s_k^*(\lambda - \phi_j - \mu)) h_0(\tau) h_0^*(\lambda)\, d\tau\, d\lambda] \\
=& \sum_i \sum_j \int \int E[\alpha_i(\tau) \alpha_j^*(\lambda)] E[\xi_i \xi_j^* s_k(\tau - \phi_i - \mu) s_k^*(\lambda - \phi_j - \mu)] h_0(\tau) h_0^*(\lambda)\, d\tau\, d\lambda \\
=& \sum_i \int \int R_\alpha(\tau, \lambda) E[|\xi_i|^2 s_k(\tau - \phi_i - \mu) s_k^*(\lambda - \phi_j - \mu)] h_0(\tau) h_0^*(\lambda)\, d\tau\, d\lambda \\
=& \int \int R_\alpha(\tau, \lambda) \sum_i \{|\xi_i|^2 E[s_k(\tau - \phi_i - \mu) s_k^*(\lambda - \phi_j - \mu)]\} h_0(\tau) h_0^*(\lambda)\, d\tau\, d\lambda \\
=& \int \int R_\alpha(\tau, \lambda) \Psi_k(\tau, \lambda) h_0(\tau) h_0^*(\lambda)\, d\tau\, d\lambda,
\end{aligned} \tag{5.30}$$

where we define

$$\Psi_k(\tau, \lambda) = \sum_i \{|\xi_i|^2 E[s_k(\tau - \phi_i - \mu) s_k^*(\lambda - \phi_j - \mu)]\}$$

Similarly,

$$\begin{aligned}
r_{k01} &= \int \int R_\alpha(\tau, \lambda) \Psi_k(\tau, \lambda) h_0(\tau) h_1^*(\lambda)\, d\tau\, d\lambda, \\
r_{k10} &= \int \int R_\alpha(\tau, \lambda) \Psi_k(\tau, \lambda) h_1(\tau) h_0^*(\lambda)\, d\tau\, d\lambda, \text{ and} \\
r_{k11} &= \int \int R_\alpha(\tau, \lambda) \Psi_k(\tau, \lambda) h_1(\tau) h_1^*(\lambda)\, d\tau\, d\lambda.
\end{aligned} \tag{5.31}$$

For the perfect synchronization situation,

$$\Psi_k(\tau, \lambda) = \sum_i \xi_i^2 s_k(\tau - \phi_i) s_k^*(\lambda - \phi_j).$$

Without the assumption of synchronization,

$$\Psi_k(\tau, \lambda) = \sum_i \xi_i^2 \int s_0(\tau - \phi_i - \mu) s_0^*(\lambda - \phi_i - \mu) p(\mu)\, d\mu,$$

where $p(\mu)$ is the density function of μ. So,

$$R_k = \begin{bmatrix} r_{k00} & r_{k01} \\ r_{k10} & r_{k11} \end{bmatrix} + \frac{N_0}{2} \begin{bmatrix} 1 & \int h_0(t) h_1(t)\, dt \\ \int h_1(t) h_0(t)\, dt & 1 \end{bmatrix} - C_k. \tag{5.32}$$

Since correlation detection is a linear operation, $\mathbf{r}$ is complex Gaussian with density function [118]

$$p(\mathbf{r}|s_k(t)) = \frac{1}{\pi^N |R_k|} \exp[-(\mathbf{r}-\bar{\mathbf{r}}_k)^H R_k^{-1}(\mathbf{r}-\bar{\mathbf{r}}_k)]. \tag{5.33}$$

Assume that both signals occur with equal prior probability. The maximum likelihood decision rule we use is

$$\begin{aligned}
&p(\mathbf{r}|s_0(t)) - p(\mathbf{r}|s_1(t)) \underset{s_1}{\overset{s_0}{\gtrless}} 0 \\
\Rightarrow &\frac{1}{\pi^2|R_0|}\exp[-(\mathbf{r}-\bar{\mathbf{r}}_0)^H R_0^{-1}(\mathbf{r}-\bar{\mathbf{r}}_0)] - \frac{1}{\pi^2|R_1|}\exp[-(\mathbf{r}-\bar{\mathbf{r}}_1)^H R_1^{-1}(\mathbf{r}-\bar{\mathbf{r}}_1)] \underset{s_1}{\overset{s_0}{\gtrless}} 0 \\
\Rightarrow &(\mathbf{r}-\bar{\mathbf{r}}_1)^H R_1^{-1}(\mathbf{r}-\bar{\mathbf{r}}_1) - (\mathbf{r}-\bar{\mathbf{r}}_0)^H R_0^{-1}(\mathbf{r}-\bar{\mathbf{r}}_0) \underset{s_1}{\overset{s_0}{\gtrless}} \ln\frac{|R_0|}{|R_1|} \\
\Rightarrow &(\mathbf{r}-\mathbf{u})^H Q(\mathbf{r}-\mathbf{u}) \underset{s_1}{\overset{s_0}{\gtrless}} \ln\frac{|R_0|}{|R_1|} + \bar{\mathbf{r}}_1^H R_1^{-1}\bar{\mathbf{r}}_1 - \bar{\mathbf{r}}_0^H R_0^{-1}\bar{\mathbf{r}}_0 - \frac{1}{4}\mathbf{w}^H Q^{-1}\mathbf{w} = \xi \\
\Rightarrow &g = \mathbf{v}^H Q \mathbf{v} \underset{s_1}{\overset{s_0}{\gtrless}} \xi,
\end{aligned} \tag{5.34}$$

where $Q = R_1^{-1} - R_0^{-1}$, $\mathbf{w} = 2(R_0^{-1}\bar{\mathbf{r}}_0 - R_1^{-1}\bar{\mathbf{r}}_1)$, $\mathbf{u} = -\frac{1}{2}Q^{-1}\mathbf{w}$, and $\mathbf{v} = \mathbf{r} - \mathbf{u}$. Let $M_0 = E[(\mathbf{v}-\bar{\mathbf{v}})(\mathbf{v}-\bar{\mathbf{v}})^H|s_0(t)]$ and $M_1 = E[(\mathbf{v}-\bar{\mathbf{v}})(\mathbf{v}-\bar{\mathbf{v}})^H|s_1(t)]$. When $s_0(t)$ is sent, we have $\mathbf{v}|_{s_0(t)} = \mathbf{r} + Q^{-1}(R_0^{-1}\bar{\mathbf{r}}_0 - R_1^{-1}\bar{\mathbf{r}}_1)$ and $\bar{\mathbf{v}}|_{s_0(t)} = \bar{\mathbf{r}}_0 + Q^{-1}(R_0^{-1}\bar{\mathbf{r}}_0 - R_1^{-1}\bar{\mathbf{r}}_1)$. Therefore, $(\mathbf{v}-\bar{\mathbf{v}})|_{s_0(t)} = (\mathbf{r}-\bar{\mathbf{r}}_0)$ which implies $M_0 = R_0$. Similarly, when $s_1(t)$ is sent, we have $M_1 = R_1$.

When we assume fast fading, $\bar{\mathbf{r}}_k = 0$, (5.15) and (5.16) can be utilized to evaluate the performance of different signaling waveforms. For $\xi \geq 0$, the bit error probability when $s_0(t)$ is sent can now be expressed as

$$\begin{aligned}
P(e|s_0(t)) &= \int_{-\infty}^{\xi} p(g|s_0(t))\, dg \\
&= F_{g|s_0(t)}(\xi) \\
&= 1 - \sum_{a_i>0} \left\{ \prod_{n=1, n\neq i}^{N} \frac{a_i}{a_i - a_n} \right\} \exp(-\xi a_i^{-1}),
\end{aligned} \tag{5.35}$$

where a_n are the eigenvalues of R_0Q. The bit error probability when $s_1(t)$ is sent,

$$\begin{aligned} P(e|s_1(t)) &= \int_{\xi}^{\infty} p(g|s_1(t))\, dg \\ &= 1 - \int_{-\infty}^{\xi} p(g|s_1(t))\, dg \\ &= 1 - F_{g|s_1(t)}(\xi) \\ &= \sum_{b_i>0} \left\{ \prod_{n=1,n\neq i}^{N} \frac{b_i}{b_i - b_n} \right\} \exp(-\xi b_i^{-1}), \end{aligned} \tag{5.36}$$

where b_n are the eigenvalues of R_1Q. Similarly, for $\xi < 0$, we have

$$P(e|s_0(t)) = \sum_{a_i<0} \left\{ \prod_{n=1,n\neq i}^{N} \frac{a_i}{a_i - a_n} \right\} \exp(-\xi a_i^{-1}) \tag{5.37}$$

and

$$P(e|s_1(t)) = 1 - \sum_{b_i<0} \left\{ \prod_{n=1,n\neq i}^{N} \frac{b_i}{b_i - b_n} \right\} \exp(-\xi b_i^{-1}). \tag{5.38}$$

The overall bit error probability is

$$P(e) = \frac{1}{2}P(e|s_0(t)) + \frac{1}{2}P(e|s_1(t)). \tag{5.39}$$

This expression will be used to characterize the performance of the waveforms created using the iterative algorithm.

Chapter 6
Implementation of the Algorithms

As mentioned in section 1.3, we are working with physically realizable signals and filters in $\mathcal{L}_2$. The difficulty in finding CPs is that solutions must be obtained by solving integral equations. These depend on the channel statistics properties. As is frequently done in numeric work in general [119], we work with sampled versions of these functions. Instead of working with the infinite dimensional functions directly, we use their approximations in finite-dimensional spaces to be able to implement these algorithms from chapter 4. As the sampling interval decreases, a solution of increasing accuracy is obtained.

6.1 Conversion of a Linear Integral Equation to Its Discrete Equivalent

A linear integral equation may be converted to a discrete equation by the following approximation. For every τ,

$$\begin{aligned} y(\tau) &= \int_a^b R(\tau,\lambda)x(\lambda)\,d\lambda \\ &= \lim_{\Delta\lambda\to 0}\sum_{m=A}^{B} R(\tau,m\Delta\lambda)x(m\Delta\lambda)\Delta\lambda, \end{aligned} \tag{6.1}$$

where $A, B \in \mathbb{N}$ such that $a \approx A\Delta\lambda$ and $b \approx B\Delta\lambda$. So, for $\Delta\lambda$ small enough, we have

$$y(\tau) \approx \sum_{m=A}^{B} R(\tau,m\Delta\lambda)x(m\Delta\lambda)\Delta\lambda. \tag{6.2}$$

Now, we can represent $y(\tau)$ in its discrete form.

$$y(\tau) = \lim_{\Delta\tau\to 0} y(n\Delta\tau), \tag{6.3}$$

where $n \in \mathbb{N}$ such that $\tau = n\Delta\tau$. Again, when $\Delta\tau$ small enough, we have

$$y(\tau) \approx y(n\Delta\tau). \tag{6.4}$$

Let $\Delta T = \min(\Delta\tau, \Delta\lambda)$, and equating (6.2) and (6.4), we obtain

$$y(n\Delta T) = \sum_{m=A}^{B} R(n\Delta T, m\Delta T)x(m\Delta T)\Delta T. \tag{6.5}$$

Define the following vectors and matrix with elements being the values of $y(n\Delta T)$, $x(m\Delta T)$, and $R(n\Delta T, m\Delta T)$, respectively:

$$\mathbf{x} = \begin{bmatrix} \vdots \\ x((n-1)\Delta T) \\ x(n\Delta T) \\ x((n+1)\Delta T) \\ \vdots \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} \vdots \\ y((n-1)\Delta T) \\ y(n\Delta T) \\ y((n+1)\Delta T) \\ \vdots \end{bmatrix}, \tag{6.6}$$

and

$$\mathbf{R} = \begin{bmatrix} & \vdots & \\ \cdots & R(m\Delta T, n\Delta T) & \cdots \\ & \vdots & \end{bmatrix}. \tag{6.7}$$

Then (6.5) is equivalent to

$$\mathbf{y} = \mathbf{R}\mathbf{x}\Delta T. \tag{6.8}$$

Thus the integral equation is approximated by a matrix equation. We will apply this conversion to the pertinent equations from chapter 4. Equations (4.32) and (4.33) have the following form

$$\begin{aligned} y(\tau) &= \int [R_\alpha(\tau,\lambda)C(\tau,\lambda) + k\,\mathrm{sinc}(2\pi f(\tau-\lambda))]y(\lambda)\,d\lambda \\ &= \int R_\alpha(\tau,\lambda)C(\tau,\lambda)y(\lambda)\,d\lambda + \int k\,\mathrm{sinc}(2\pi f(\tau-\lambda))y(\lambda)\,d\lambda. \end{aligned} \tag{6.9}$$

In (6.9), we have the discrete equivalent of $\int R_\alpha(\tau,\lambda)C(\tau,\lambda)y(\lambda)\,d\lambda$ being $\mathbf{P}\mathbf{y}$, where

$$\mathbf{P} = \begin{bmatrix} & \vdots & \\ \cdots & R_\alpha(m\Delta T, n\Delta T)C(m\Delta T, n\Delta T) & \cdots \\ & \vdots & \end{bmatrix} = \mathbf{R}_\alpha \circ \mathbf{C}. \tag{6.10}$$

In (6.10), $\mathbf{R}_\alpha$ is the matrix representations of $\mathbf{R}_\alpha(\tau,\lambda)$, $\mathbf{C}$ is the matrix representation of $C(\tau,\lambda)$, and $\circ$ represents the Hadamard (element by element) product. So, (6.9) can now be represented in discrete form as

$$\mathbf{y} = [\mathbf{R}_\alpha \circ \mathbf{C} + k\mathbf{W}]\mathbf{y}, \tag{6.11}$$

where $\mathbf{W}$ is the matrix representation of $2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))$.

6.2 Discrete Version of Algorithm 2 and 3

Let $\mathbf{R}_\alpha$, Ψ,Θ, $\mathbf{W}$, and $\mathbf{S}$ be the matrix representation of $R_\alpha(\tau,\lambda)$, $\Psi(\tau,\lambda)$, $\Theta(\tau,\lambda)$, $2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))$, and $\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$, respectively. Algorithm 2 can be presented in sampled form.

Algorithm 4 *1. Randomly select an initial vector* $\mathbf{s}_0$ *of length N.*

2. Set $\mathbf{h}_0$ *to make*

$$\gamma_1\bar{\mathbf{h}}_0 = [(1-\beta)\mathbf{R}_\alpha \circ \Psi + \beta\mathbf{W}]\bar{\mathbf{h}}_0, \tag{6.12}$$

where γ_1 *is the largest eigenvalue in magnitude of the kernel* $[(1-\beta)\mathbf{R}_\alpha\Psi + \beta\mathbf{W}]$, *while* $\bar{\mathbf{h}}_0$ *is the corresponding eigenvector.*

3. Set $\mathbf{s}_0$ *so that for* $\mathbf{h}_0$ *from the last step,*

$$\gamma_0\bar{\mathbf{s}}_0 = [(1-\beta)\mathbf{R}_\alpha \circ \Theta + \beta\mathbf{W}]\bar{\mathbf{s}}_0, \tag{6.13}$$

where γ_0 *being the largest eigenvalue in magnitude of the kernel* $[(1-\beta)\mathbf{R}_\alpha\Theta + \beta W]$, *and* $\bar{\mathbf{s}}_0$ *being the corresponding eigenvector.*

4. Repeat steps 2, 3, and 4 a sufficient number of times.

Algorithm 3 can be presented in discrete sampled form.

Algorithm 5 *1. Randomly select initial functions* $\mathbf{s}_0$ *and* $\mathbf{s}_1$.

2. Set $\mathbf{h}_{00}$ *so that*

$$\gamma_{00}\bar{\mathbf{h}}_{00} = [(1-\beta)\mathbf{R}_\alpha \circ \Psi_0 + \beta\mathbf{W}]\bar{\mathbf{h}}_{00}, \tag{6.14}$$

γ_{00} *is the largest eigenvalue of the kernel* $[(1-\beta)\mathbf{R}_\alpha \circ \Psi_0 + \beta\mathbf{W}]$, *while* $\bar{\mathbf{h}}_{00}$ *is the corresponding eigenfunction.*

3. Set $\mathbf{h}_{01}$ *so that*

$$\xi_{01}\bar{\mathbf{h}}_{01} = [(1-\beta)\mathbf{R}_\alpha \circ \Psi_0 + \beta\mathbf{S}]\bar{\mathbf{h}}_{01}, \tag{6.15}$$

ξ_{11} *is the smallest eigenvalue of the kernel* $[(1-\beta)\mathbf{R}_\alpha \circ \Psi_0 + \beta\mathbf{S}]$, *while* $\bar{\mathbf{h}}_{01}$ *is the corresponding eigenfunction.*

4. *Set* $\mathbf{h}_{11}$ *so that*

$$\gamma_{11}\bar{\mathbf{h}}_{11} = [(1-\beta)\mathbf{R}_\alpha \circ \Psi_1 + \beta\mathbf{W}]\bar{\mathbf{h}}_{10}, \tag{6.16}$$

γ_{11} is the largest eigenvalue of the kernel $[(1-\beta)\mathbf{R}_\alpha \circ \Psi_1 + \beta\mathbf{W}]$, while $\bar{\mathbf{h}}_{11}$ is the corresponding eigenfunction.

5. *Set* $\mathbf{h}_{10}$ *so that*

$$\xi_{10}\bar{\mathbf{h}}_{10} = [(1-\beta)\mathbf{R}_\alpha \circ \Psi_1 + \beta\mathbf{S}]\bar{\mathbf{h}}_{10}, \tag{6.17}$$

ξ_{10} is the smallest eigenvalue of the kernel $[(1-\beta)\mathbf{R}_\alpha \circ \Psi_1 + \beta\mathbf{S}]$, while $\bar{\mathbf{h}}_{10}$ is the corresponding eigenfunction.

6. *Find the average, i.e.,*

$$\mathbf{h}_0 = \mathbf{h}_{00} + \mathbf{h}_{10}, \tag{6.18}$$

and

$$\mathbf{h}_1 = \mathbf{h}_{01} + \mathbf{h}_{11}. \tag{6.19}$$

7. *Normalize* $\mathbf{h}_0$ *and* $\mathbf{h}_1$.

8. *Set* $\mathbf{s}_{00}$ *so that*

$$\gamma_{00}\bar{\mathbf{s}}_{00} = [(1-\beta)\mathbf{R}_\alpha \circ \Theta_0 + \beta\mathbf{W}]\bar{\mathbf{s}}_{00}, \tag{6.20}$$

γ_{00} is the largest eigenvalue of the kernel $[(1-\beta)\mathbf{R}_\alpha \circ \Theta_0 + \beta\mathbf{W}]$, while $\bar{\mathbf{s}}_{00}$ is the corresponding eigenfunction.

9. *Set* $\mathbf{s}_{01}$ *so that*

$$\xi_{01}\bar{\mathbf{s}}_{01} = [(1-\beta)\mathbf{R}_\alpha \circ \Theta_0 + \beta\mathbf{S}]\bar{\mathbf{s}}_{01}, \tag{6.21}$$

ξ_{11} is the smallest eigenvalue of the kernel $[(1-\beta)\mathbf{R}_\alpha \circ \Theta_0 + \beta\mathbf{S}]$, while $\bar{\mathbf{s}}_{01}$ is the corresponding eigenfunction.

10. *Set* $\mathbf{s}_{11}$ *so that*

$$\gamma_{11}\bar{\mathbf{s}}_{11} = [(1-\beta)\mathbf{R}_\alpha \circ \Theta_1 + \beta\mathbf{W}]\bar{\mathbf{s}}_{10}, \tag{6.22}$$

γ_{11} is the largest eigenvalue of the kernel $[(1-\beta)\mathbf{R}_\alpha \circ \Theta_1 + \beta\mathbf{W}]$, while $\bar{\mathbf{s}}_{11}$ is the corresponding eigenfunction.

11. Set $\mathbf{s}_{10}$ *so that*

$$\xi_{10}\bar{\mathbf{s}}_{10} = [(1-\beta)\mathbf{R}_\alpha \circ \Theta_1 + \beta\mathbf{S}]\bar{\mathbf{s}}_{10}, \tag{6.23}$$

ξ_{10} is the smallest eigenvalue of the kernel $[(1-\beta)\mathbf{R}_\alpha \circ \Theta_1 + \beta\mathbf{S}]$, while $\bar{\mathbf{s}}_{10}$ is the corresponding eigenfunction.

12. Find the average, i.e.,

$$\mathbf{s}_0 = \mathbf{s}_{00} + \mathbf{s}_{10}, \tag{6.24}$$

and

$$\mathbf{s}_1 = \mathbf{s}_{01} + \mathbf{s}_{11}. \tag{6.25}$$

13. Normalize $\mathbf{s}_0$ *and* $\mathbf{s}_1$.

14. Repeat steps 2 to 13 a sufficient number of times. Save these CCPs with a large SNR.

As $R_\alpha(\tau, \lambda)$ is an autocorrelation function, while $2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))$ and $\mathcal{I}[2f\,\mathrm{sinc}(2\pi f(\tau-\lambda))]$ are the inverse Fourier transform of non-negative functions, all these kernels are positive semi-definite. So, $\mathbf{R}_\alpha$, $\mathbf{W}$, and $\mathbf{S}$ are required to be positive semi-definite, too. Additionally, since they are Toeplitz matrices, the requirement of positive semi-definite implies the discrete-time Fourier transforms of the sequences of their elements have to be non-negative [120]. In general, a sampled sequence is not necessary non-negative. This is because when we perform discrete sampling from a continuous function, g, and present it as a finite dimensional vector, we introduce two artifacts. These are the windowing effect as we truncate the function to a finite length, and the aliasing effect if we do not sample fast enough. Aliasing can be controlled fairly well since most physical autocorrelation functions are low-pass. When we perform any kind of windowing, we equivalently convolve the Fourier transform of g with the Fourier transform of the windowing function. By performing this convolution, we introduce negative values to the Fourier transform of the resulting function. The effect of introducing negative values to a positive sequence can only be reduced, but cannot be completely eliminated, if we use a windowing function that has a long support.

The sampled sequence can be made to have a positive transform again by running it through a composite mapping algorithm [121]. However, this usually introduce a delta function at $g(0)$, which causes a deviation of the sample vector from the original function. Fig. 6.1 illustrates this deviation and the new peak to $g(0)$.

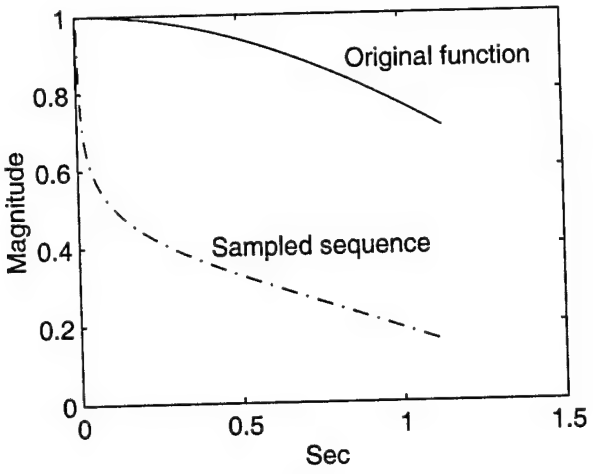


Fig. 6.1: The illustration of windowing effect.

6.3 Multipath Fading Channel Model

We use the autocorrelation function of the land-mobile model (as shown in Table 2.1) $\mathbf{R}_\alpha(\tau) = J_0(2\pi B_d \tau)$ to represent the fast fading aspect, and the European Global System for Mobile Communications (GSM) typical-urban (TU) and hilly-terrain (HT) models to represent the multipath aspect of our study [96]. Table 6.1 presents the amplitude and delay profiles of the simpler 6-path TU and HT multipath channels.

Table 6.1: Delay and Amplitude Profile of GSM Channels.

	TU model		HT model	
Path #	Delay(μs)	Power	Delay(μs)	Power
1	0.000	1.000	0.000	1.000
2	0.813	0.669	0.813	0.251
3	1.626	0.448	1.626	0.060
4	2.439	0.300	15.447	2.258
5	3.252	0.200	16.260	0.177
6	4.056	0.134	17.073	0.122

6.4 Numerical Methods

The most prominent challenge we face in computing the numerical solution to (6.12), (6.13) and the similar equations in Algorithm 5, is the size of the matrices. We need to solve for the eigenvectors corresponding to the largest eigenvalue and the smallest eigenvalue of these matrices. A large matrix size exists because we need a large sample size to take off the original function before truncation. For example, we need a span of at least 0 to 3 seconds for $J_0(\cdot)$. Only then the sampled sequence resemble the original sequence loosely. This is shown in Fig. 6.2. This is similar to Fig. 6.1, which is shown at a different time scale. We can observe the same delta function at time 0 but with a smaller magnitude. Also shown in Fig. 6.2, half of the the sample sequence is set to zero in order to get a proper autocorrelation matrix. So, the actual approximation is done for the range of t from 0 to 3 seconds. Now, some simple arithmetic can show us the magnitude of the sample

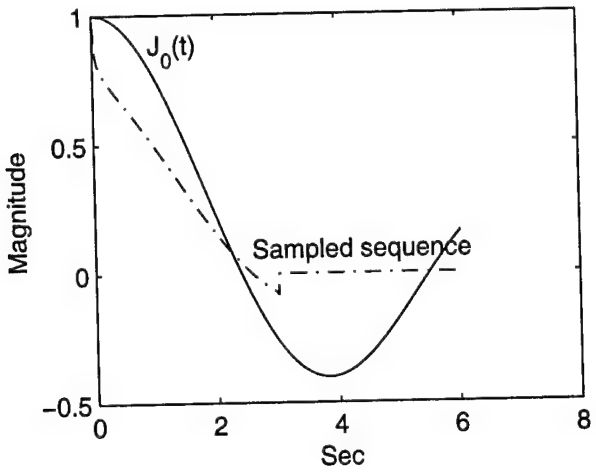


Fig. 6.2: Approximation of $J_0(t)$ with a positive sequence.

size we need. From Table 6.1, we can see that the path delays are multiples of $0.813\mu s$. It is reasonable for us to use one-tenth of the unit of the path delays as our sampling time, i.e., $\delta T = 8.13 \times 10^{-8} s$, where δT is the sampling interval. So, We have

$$2\pi B_d \delta T n = 3 \Rightarrow B_d n = 5872876.129, \tag{6.26}$$

where n is the sample number corresponding to $t = 3$ seconds, B_d is the Doppler spread and is defined as $B_d = \frac{v}{\lambda}$, where v is the travel speed, and λ is the wave-length of the

carrier waveform. Suppose we are walking at a speed of 2 meters per second while talking on a cellular phone with a frequency of 900MHz. Then, $B_d = 6$Hz. This will mean we need a sample size of 978,813! This number corresponds to a $978,813 \times 978,813$ matrix whose eigenvalues would have to be computed. We may reduce the number by changing the setting. For example, if we are traveling in a fast car at 80 miles per hour while talking on a cellular phone with a frequency of 1800MHz, we have $B_d = 216$Hz. Furthermore, if we use $\delta T = 1.626 \times 10^{-7}$s, we need a sample size of 13,595. Although this is a much more implementable figure, we still need to improve the method of numerical computation so that the algorithm may be implemented on today's computers. Mainly, we want to reduce the memory requirement and eliminate unnecessary computations. For example, the eigenvectors corresponding to the largest eigenvalues can be calculated by using the power method. As the kernel of the linear equation is symmetric, the eigenvectors corresponding to the smallest eigenvalues can be calculated by using the conjugate gradient method [122]. So, we can avoid having to compute all the eigenvectors.

The power method is an iterative method that repeatedly multiplies an initial random vector by the matrix for which we intend to find the eigenvector corresponding to the largest eigenvalue in magnitude (which may be either positive or negative). By this process, the product of the matrix-vector multiplication converges to the eigenvector with largest eigenvalue in magnitude. The method requires that there be a single eigenvalue of largest magnitude. The convergence rate of the power method is highly dependent on the relative magnitudes of the eigenvalues. In the extreme case that all the eigenvalues have the same magnitude, the power method will not converge.

The conjugate gradient method is an iterative method that solves an equation of the form

$$\mathbf{Ax} = \mathbf{b}, \tag{6.27}$$

which is recommended by Demmel as the algorithm of choice for systems involving symmetric positive definite matrices [122]. What this method does is actually find $\mathbf{x}$ that minimizes the function

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{xAx} - \mathbf{bx}. \tag{6.28}$$

When the above function is minimized, the gradient

$$\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b} \tag{6.29}$$

is zero. The conjugate gradient method itself does not solve the eigenvalue problem. By iteratively applying the conjugate gradient method to an initial random vector, we have the inverse power method. The process converges to the eigenvector corresponding to the smallest eigenvalue in magnitude. The convergence rate of the conjugate gradient method again depends on the relative magnitude of the eigenvalues of the matrix. The closer the small eigenvalues are to each other, the slower the convergence.

If we examine the expression of the equations in Algorithms 4 and 5 more closely, they are of the following form

$$\gamma\bar{\mathbf{u}} = [\beta R \circ (\mathbf{v}\mathbf{v}^H) + (1-\beta)Q]\bar{\mathbf{u}}, \tag{6.30}$$

for some vector $\mathbf{v}$. Since R and Q are Toeplitz Hermitian matrices, the entire computation can be done with vector manipulations. This can drastically reduce the storage used. Every $n \times n$ matrix can be represented using an $n \times 1$ vector.

In fact, both the power method and the conjugate gradient method only require being able to compute matrix-vector multiplications, i.e., we need to be able to calculate

$$[R \circ (\mathbf{v}\mathbf{v}^H) + Q]\mathbf{x} = \mathbf{y}, \tag{6.31}$$

assuming that β and $(1-\beta)$ are absorbed into R and Q, respectively. Let $\mathbf{r}$ be the coefficients of R, and $\mathbf{q}$ the coefficients of Q. For (6.31), each term of $\mathbf{y}$ can be evaluated as

$$y_i = \sum_j [r_{(j-i)} v_i v_j^* + q_{(j-i)}] x_j \tag{6.32}$$

The power method and the conjugate gradient method in cooperation with the above method of matrix-vector multiplication are used to evaluate Algorithm 4. A result is shown in Fig. 6.3 with $8192 = 2^{13}$ sample points. Fig. 6.3, like many other plots in chapter 7, is arranged in groups of four. The top left-hand plot depicts the initial random

signal while the top right-hand plot depicts the change of the eigenvalue through the iterations. The bottom left-hand plot depicts the final **s** while the bottom right-hand plot depicts the final **h**. There are two curves in the plot of eigenvalues. One of the curves is obtained from the iteration of the **s** part while the other curve is obtained from the iteration of the **h** part. Mostly, these two curves are close to each other and are nearly indistinguishable. However, when the algorithm does not converge, these curves can differ widely. For analytical purposes, there is no essential difference in the properties of these two sequences of eigenvalues. Therefore, these curves need no specific labels.

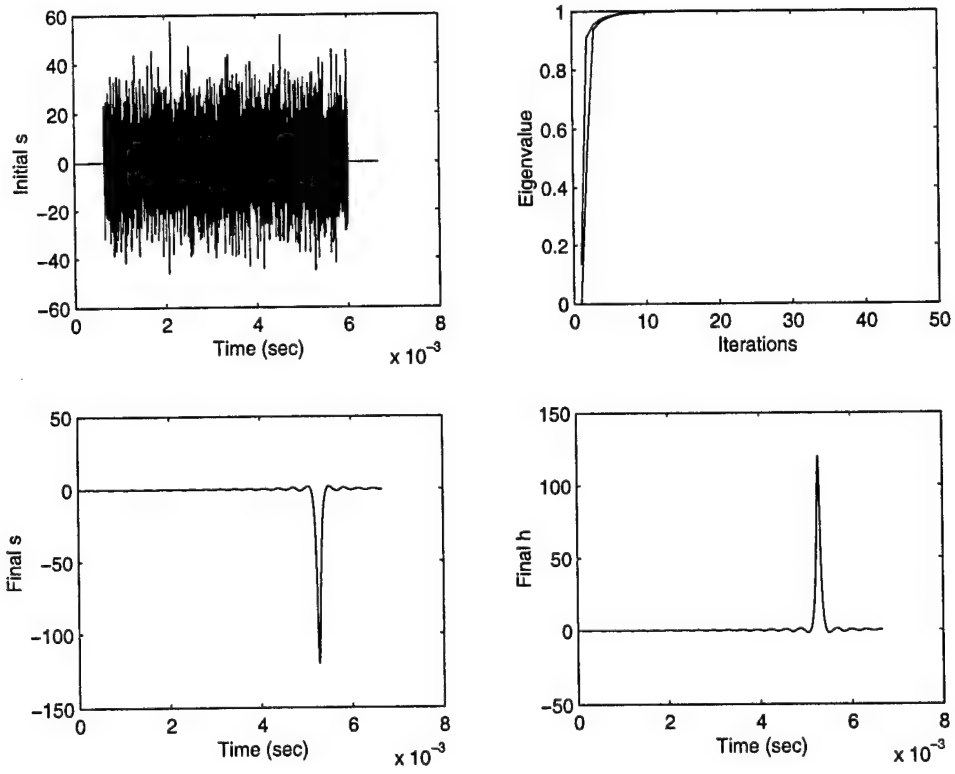


Fig. 6.3: The illustration of the result from Algorithm 4 with 8192 data points.

Although our program is capable of generating the above results within 10 iterations, for any sample size that is larger than around 2^{13}, our program will require significantly more time to run. Because the use of a smaller sample size merely implies a nonrealistic B_d, which does not compromise the characteristic of the CP that relates to the channel properties, all of our experiments were run with sample size 2^{11}. We will use the HT and

TU channel profiles (as shown in Table 6.1) as the multipath model. In order to observe the fast fading characteristic, the channel needs to vary fast enough. At the same time we prefer to have a positive sequence that closely approximates the original channel autocorrelation function. These together require a large B_d. In fact, we use $B_d = 105,000$. We can obtain a B_d of $105,000$ when we travel at $4,088$m/s while using a carrier frequency of 7.7GHz. The parameter for frequency containment is set to be $f = 6,250$Hz, $f = 12,500$Hz, and $f = 25,000$Hz.

Chapter 7
Results and Observations

Some experimental results from Algorithms 4 and 5 are presented in this chapter. The multipath models as presented in section 6.3 are used. The sequence of the largest eigenvalues obtained through iterations of the algorithms is used as the measure of both the performance of the CPs and the convergence of the algorithms. We point out that for a single path channel, the highest value of the largest eigenvalue from the algorithms is 1. This value serves as a reference standard.

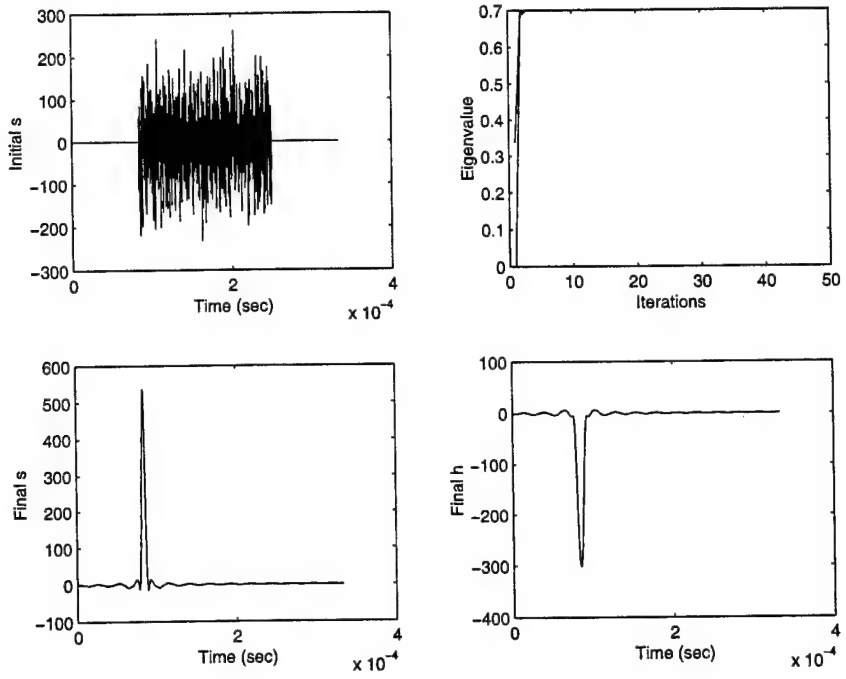
The results are grouped by aspects of interest and are presented in the following subsections: different initial vectors, frequency containment, the multipath characteristics, the effect of synchronization uncertainty, and the performance of the signal sets.

7.1 Different Initial Vectors

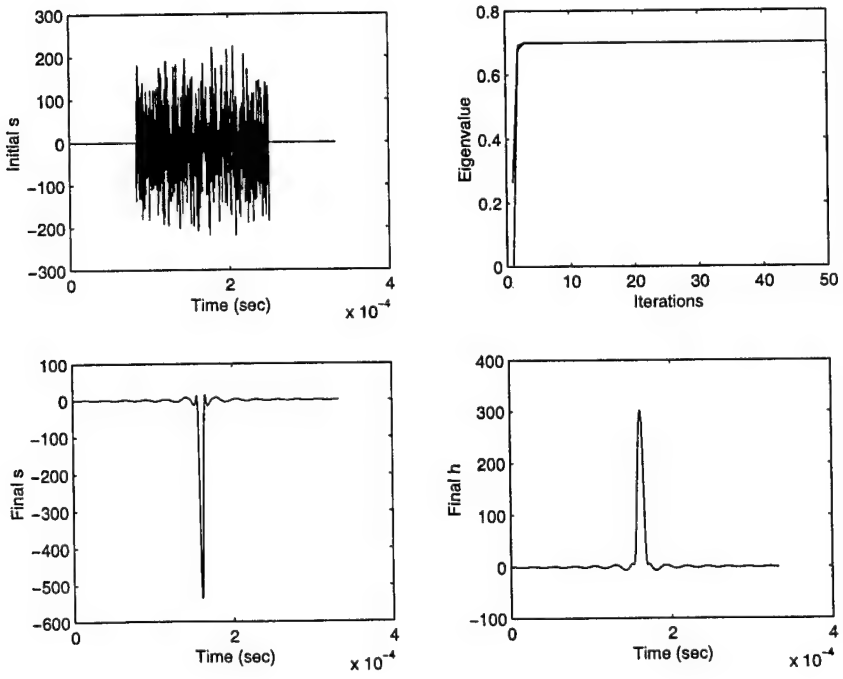
As we are not certain about the convergence of Algorithm 4, it is assuring to observe its behavior for different initial vectors. The examples presented here are with TU multipath, $\beta = 0.3$, $f = 12500$Hz, and uniformly distributed synchronization delay with distribution of $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec. As expected, Figs. 7.1 and 7.2 show that for different initial vectors, the resulting CPs are very similar to each other up to a shift in time. The largest eigenvalues converge to 0.6994 in all cases.

7.2 Frequency Containment

All plots in this section use $f = 12,500$Hz. They are divided into four groups: Group 1, as shown in Figs. 7.3 and 7.4, comprises single path fading without synchronization uncertainty; Group 2, as shown in Figs. 7.5 and 7.6, comprises single path fading with synchronization uncertainty whose delay is Gaussian-distributed ($\sigma = 1.124$); Group 3, as shown in Figs. 7.7 and 7.8, comprises multipath channel (HT) without synchronization uncertainty; and Group 4, as shown in Figs. 7.9 and 7.10, comprises multipath channel

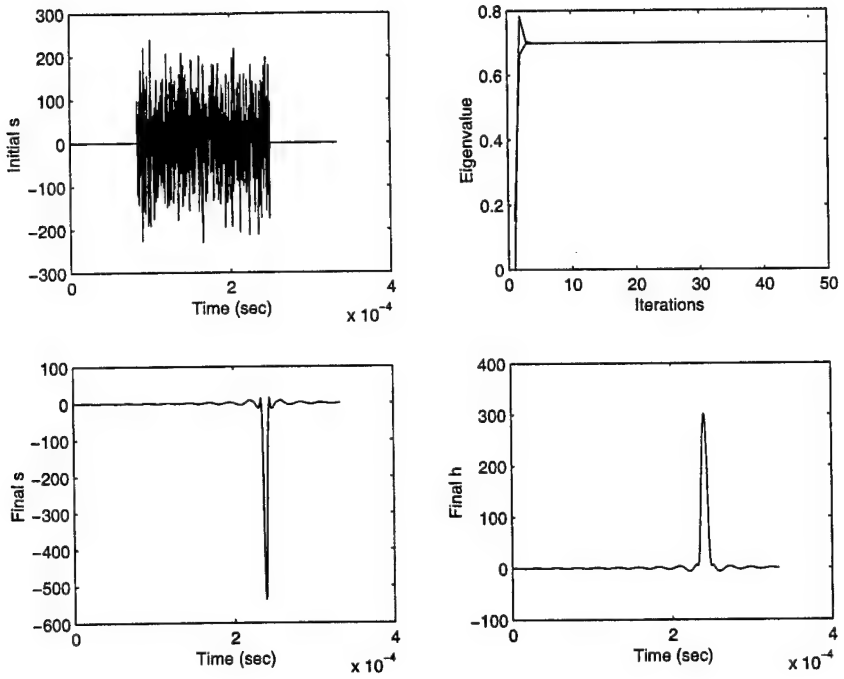


(a) First run.

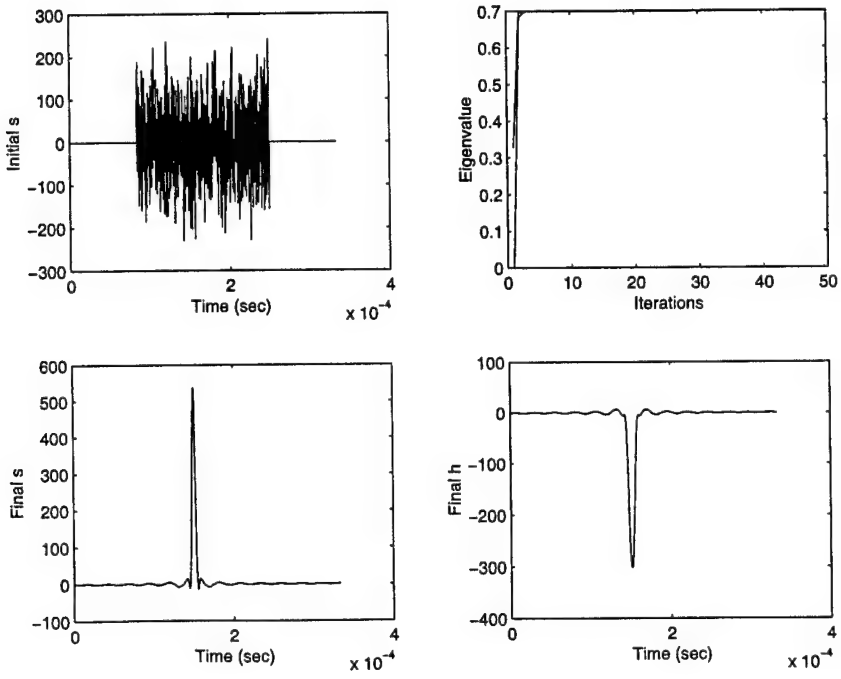


(b) Second run.

Fig. 7.1: Illustration of the effect of different initial vectors I.

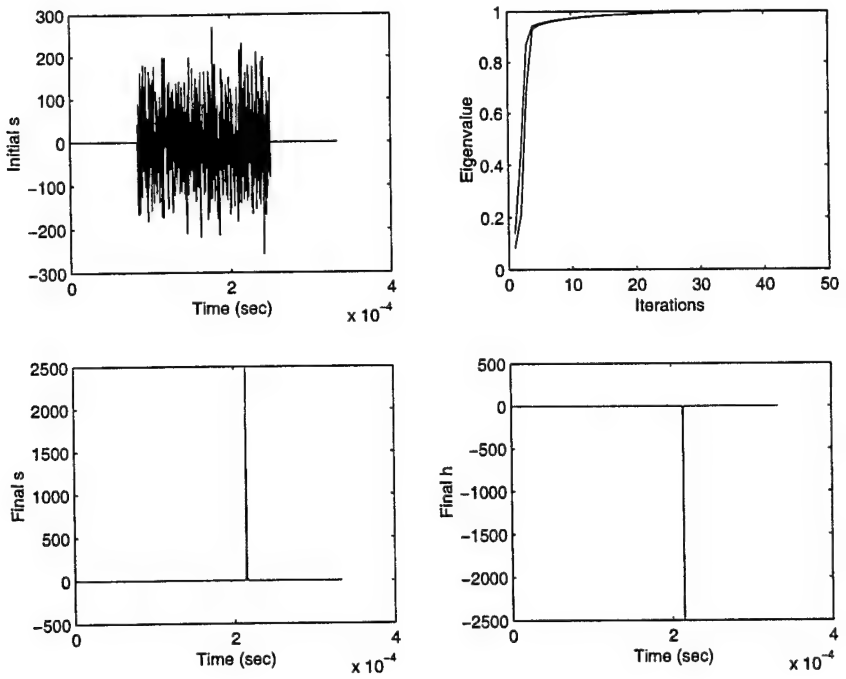


(a) Third run.

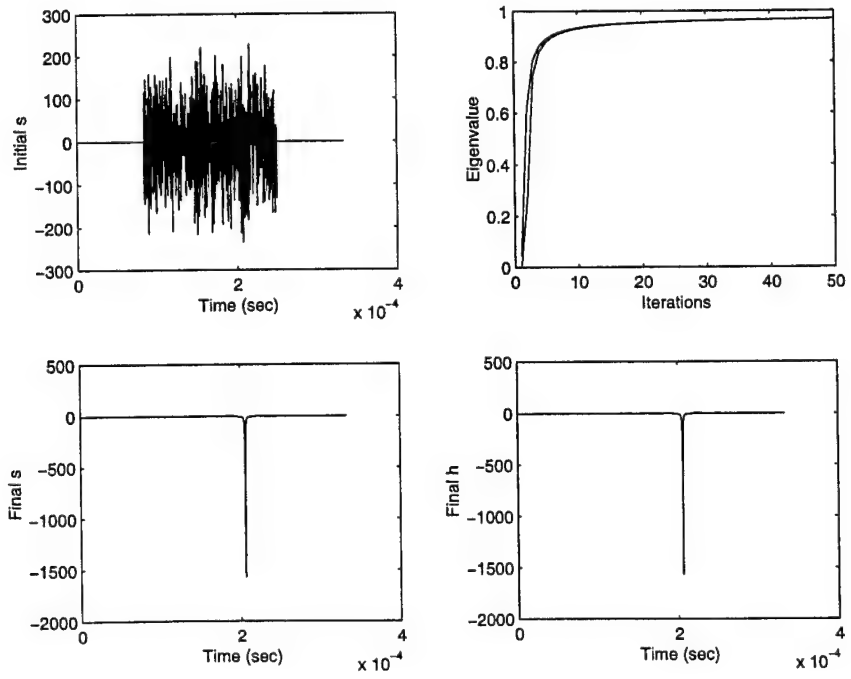


(b) Forth run.

Fig. 7.2: Illustration of the effect of different initial vectors II.

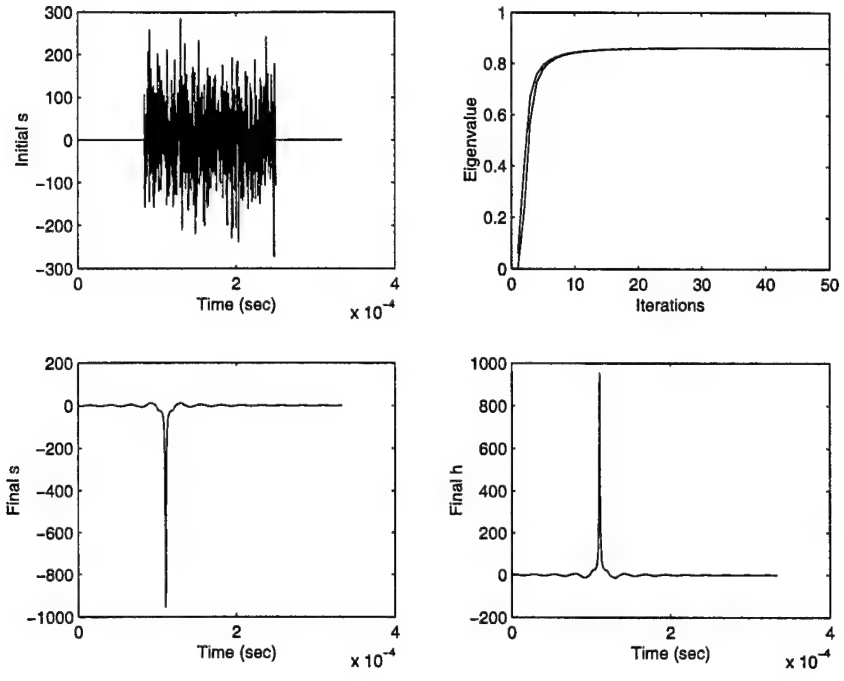


(a) $\beta = 0.1$

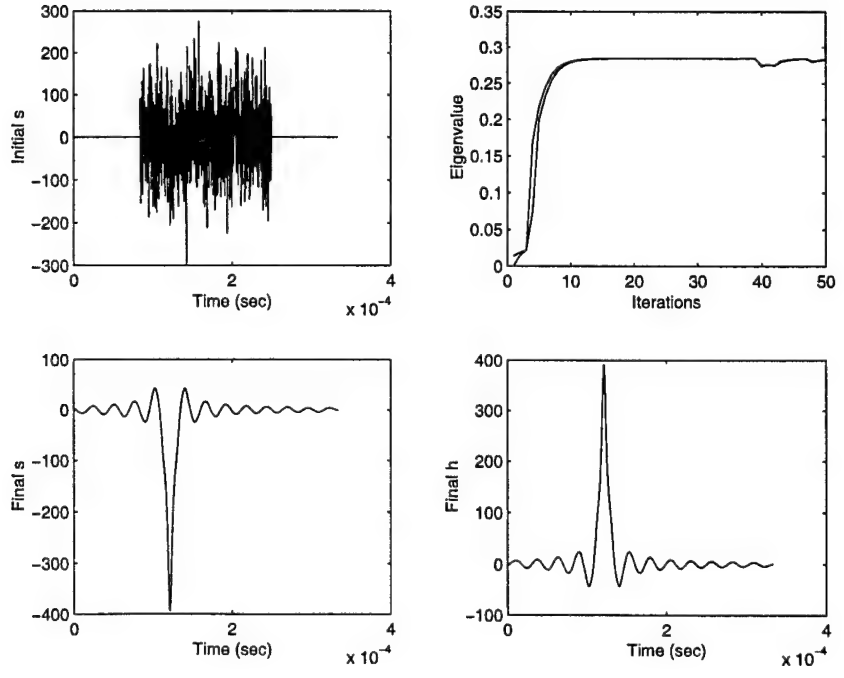


(b) $\beta = 0.3$

Fig. 7.3: Single path, perfect synchronization, $\beta = 0.1$ and 0.3.

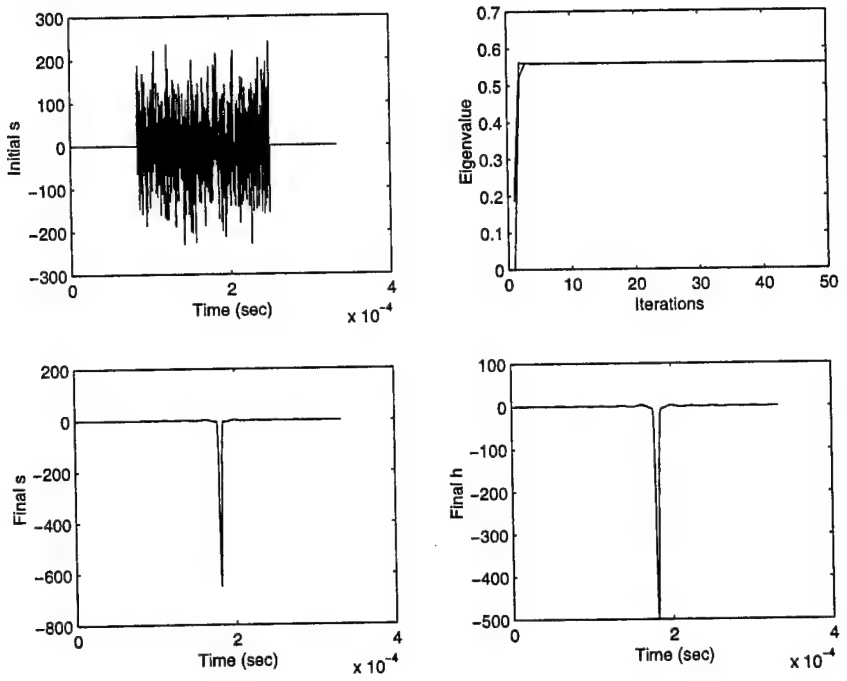


(a) $\beta = 0.5$

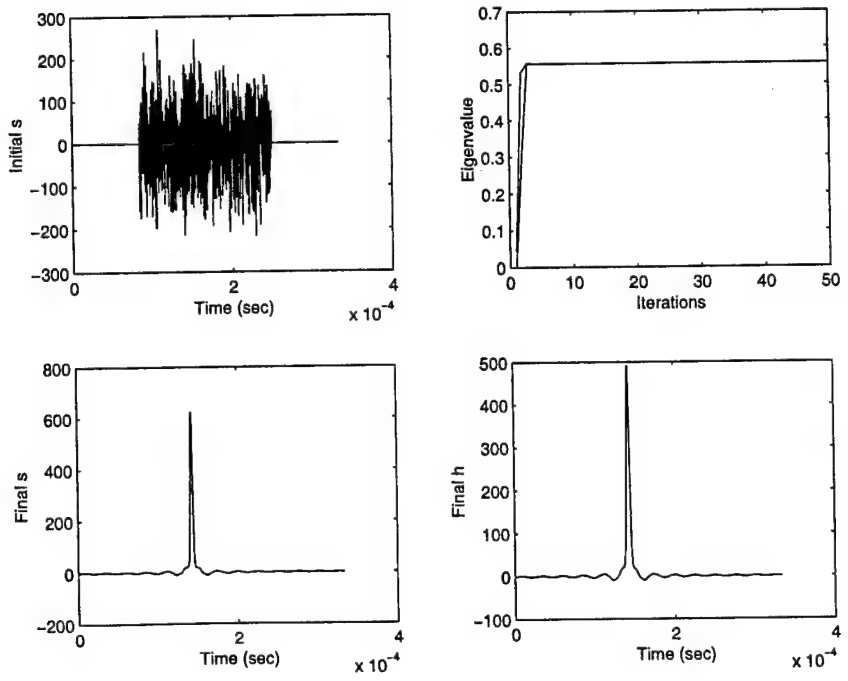


(b) $\beta = 0.7$

Fig. 7.4: Single path, perfect synchronization, $\beta = 0.5$ and 0.7.

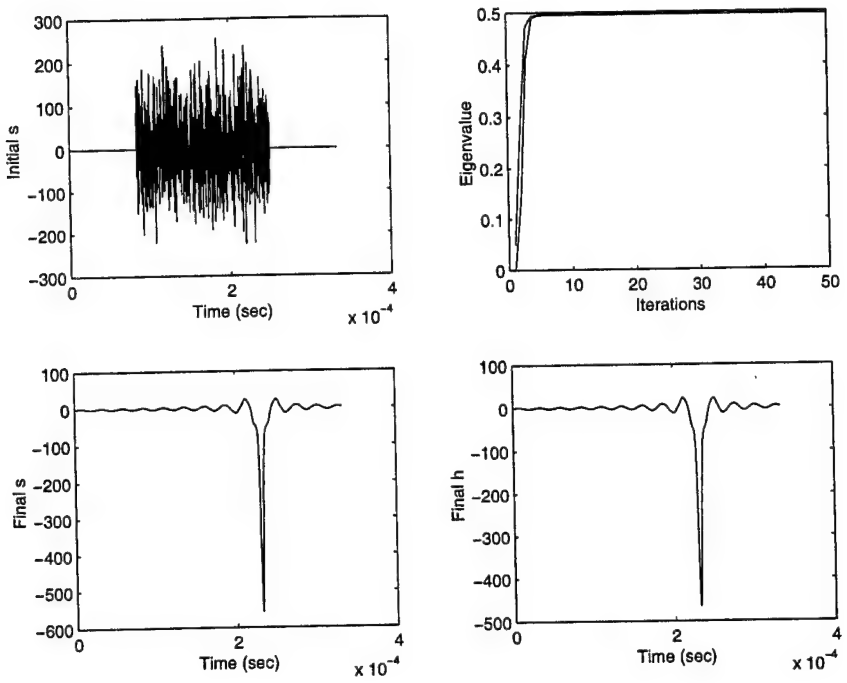


(a) $\beta = 0.1$

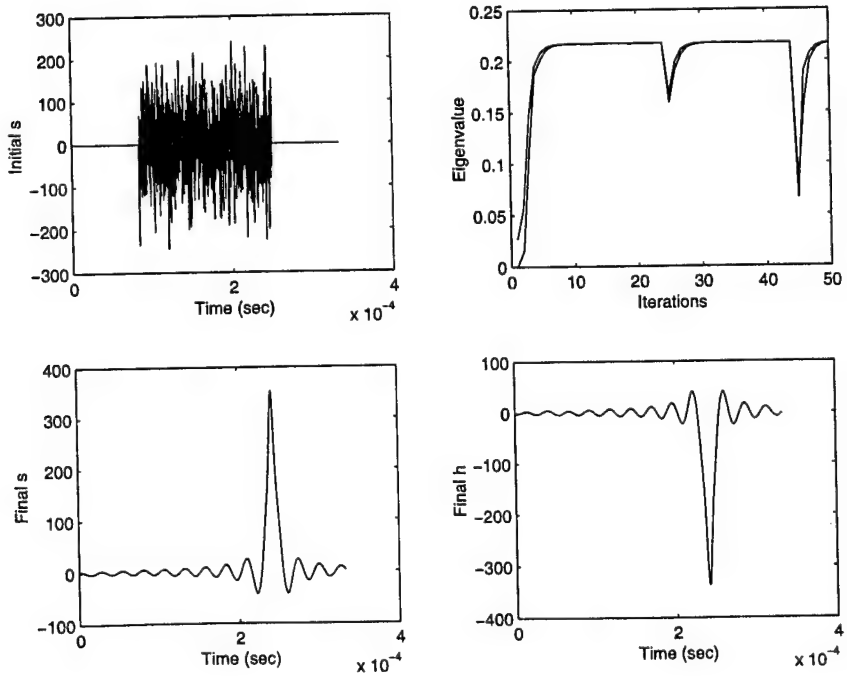


(b) $\beta = 0.3$

Fig. 7.5: Single path, Gaussian distribution with $\sigma = 1.124$, $\beta = 0.1$ and 0.3.

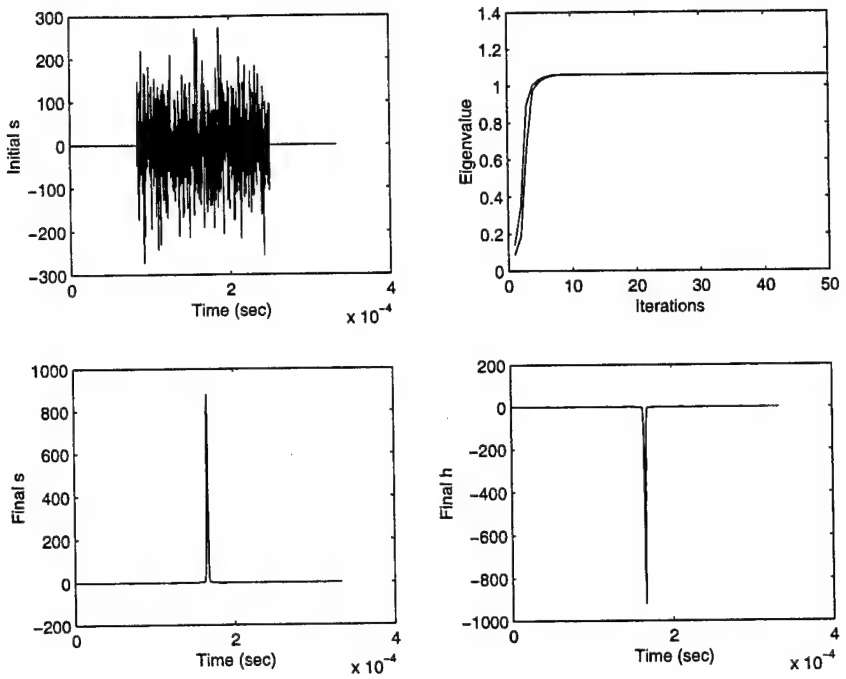


(a) $\beta = 0.5$

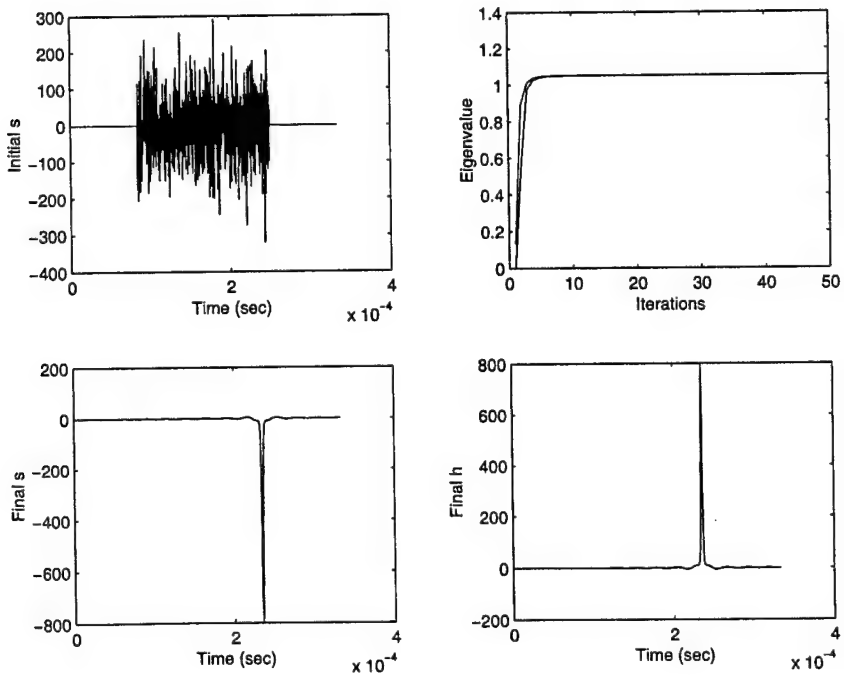


(b) $\beta = 0.7$

Fig. 7.6: Single path, Gaussian distribution with $\sigma = 1.124$, $\beta = 0.5$ and 0.7.

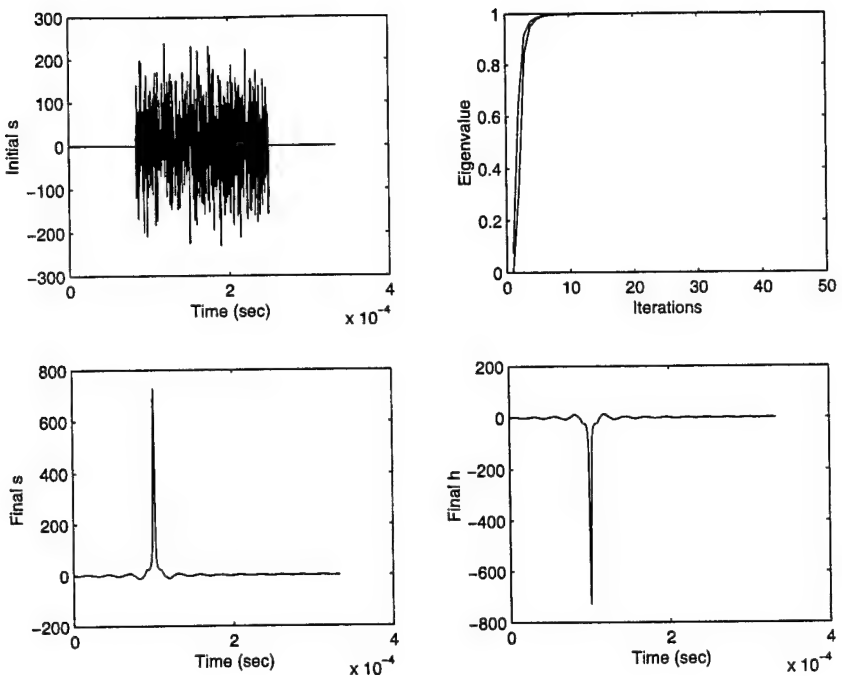


(a) $\beta = 0.1$

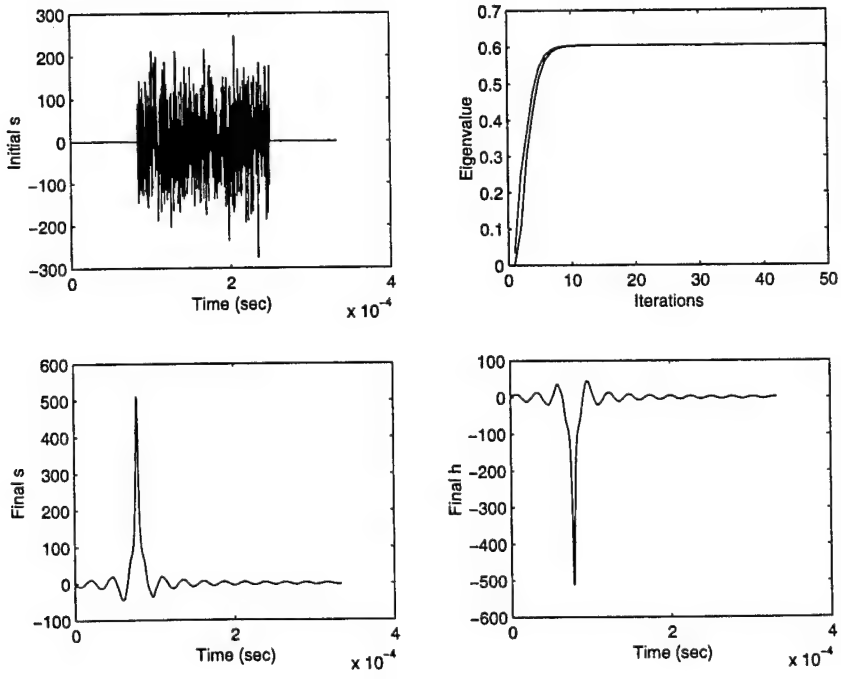


(b) $\beta = 0.3$

Fig. 7.7: HT, perfect synchronization, $\beta = 0.1$ and 0.3.

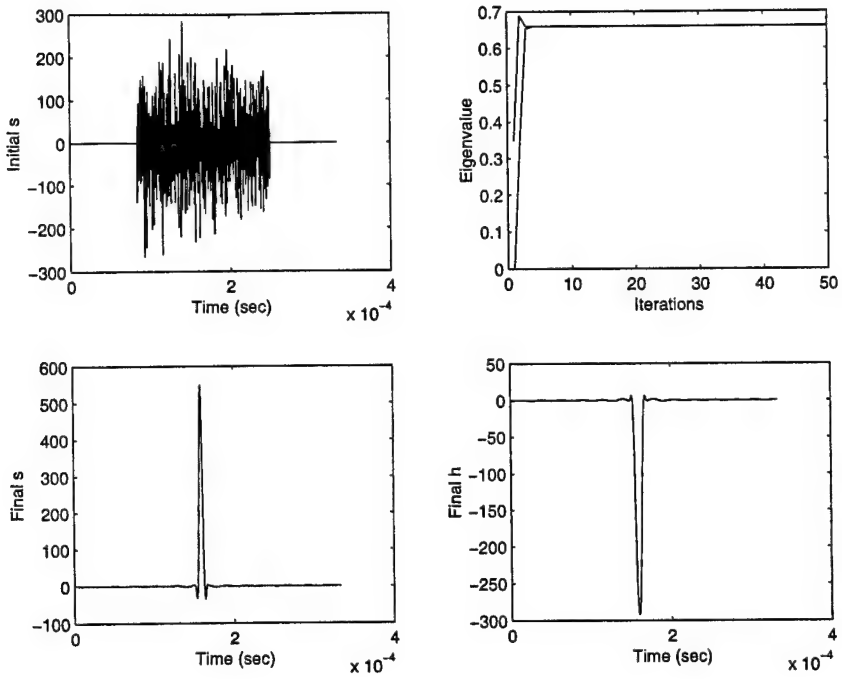


(a) $\beta = 0.5$

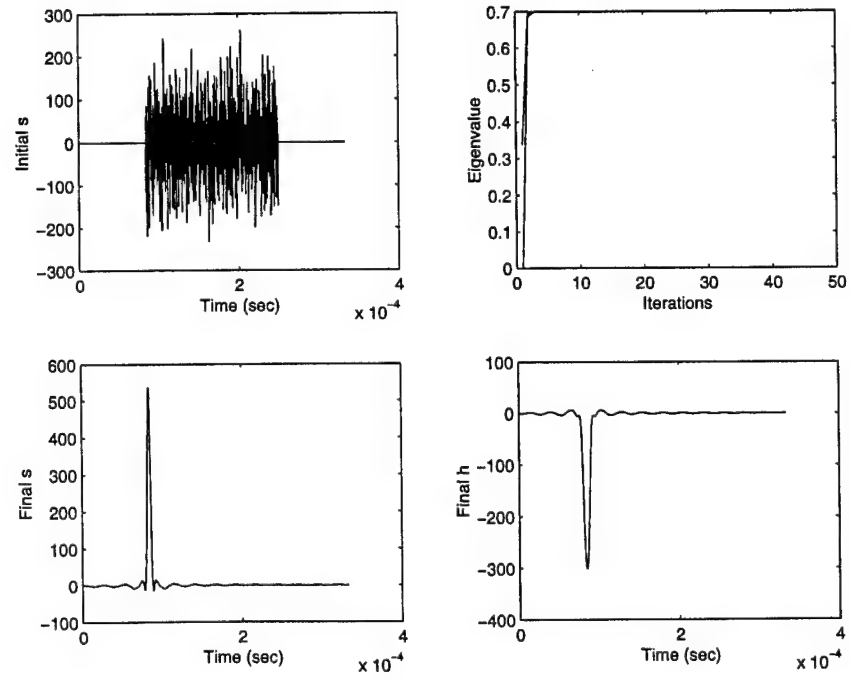


(b) $\beta = 0.7$

Fig. 7.8: HT, perfect synchronization, $\beta = 0.5$ and 0.7.

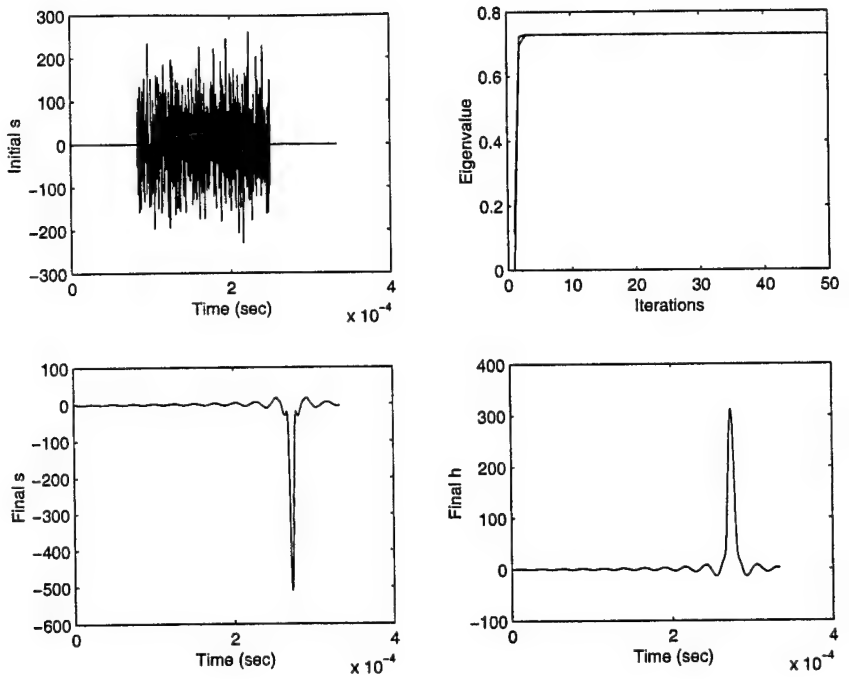


(a) $\beta = 0.1$

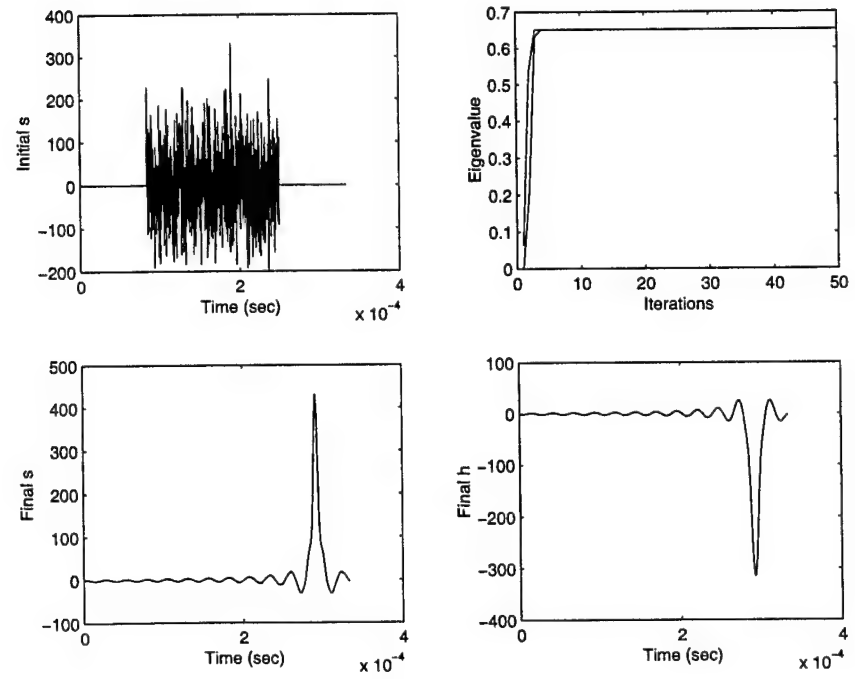


(b) $\beta = 0.3$

Fig. 7.9: TU, uniformly distributed, $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec, $\beta = 0.1$ and 0.3.



(a) $\beta = 0.5$



(b) $\beta = 0.7$

Fig. 7.10: TU, uniformly distributed, $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec, $\beta = 0.5$ and 0.7.

(TU) with uniformly distributed synchronization delay, $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec. We can observe from all these plots that the signal and filter functions get smoother with a decrease in the largest eigenvalue as β increases. We can observe in Figs. 7.4(b) and 7.6(b) some numerical errors because the power method does not converge fast enough.

7.3 Effect of Multipath Channels

The most interesting observation for multipath channels is that multiple paths may increase the SNR, compared with single paths. All of the plots in Figs. 7.11 and 7.12 have $\beta = 0$. Fig. 7.11(a) shows a plot with single path fading, which has its largest eigenvalue converged to 1. The multipath case in Fig. 7.11(b) has its largest eigenvalue converged to 1.6496.

We also observe that the signals adapt to the channel properties. The peaks shown in Fig. 7.11(b) are separated exactly by channel path delays, as expected. When there are synchronization uncertainties, the peaky nature of the pulse disappears as shown in Figs. 7.12(a) and (b). However, the largest eigenvalues still have higher values compared to that of single path settings.

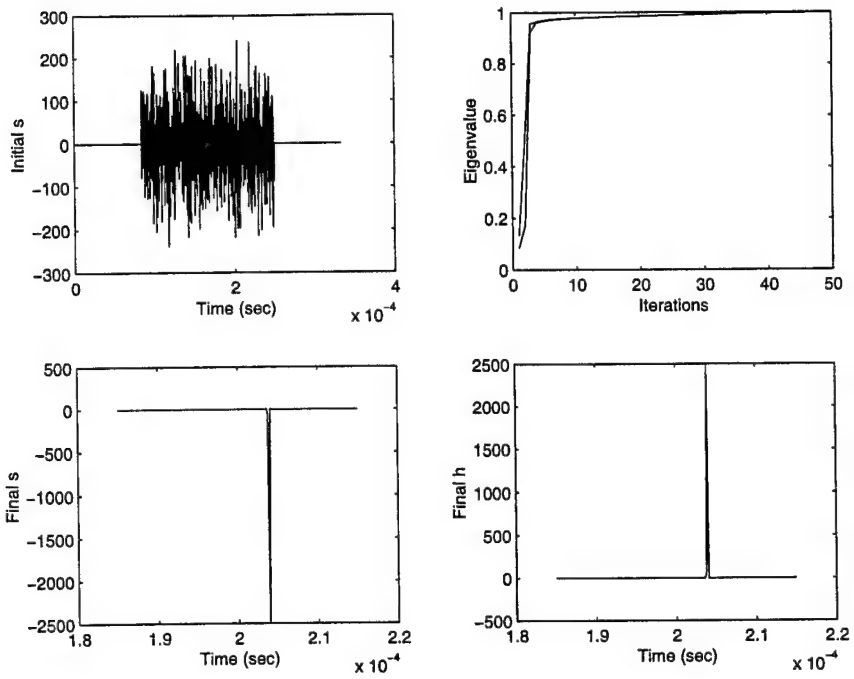
7.4 Effect of Synchronization Uncertainty

7.4.1 Fast Fading

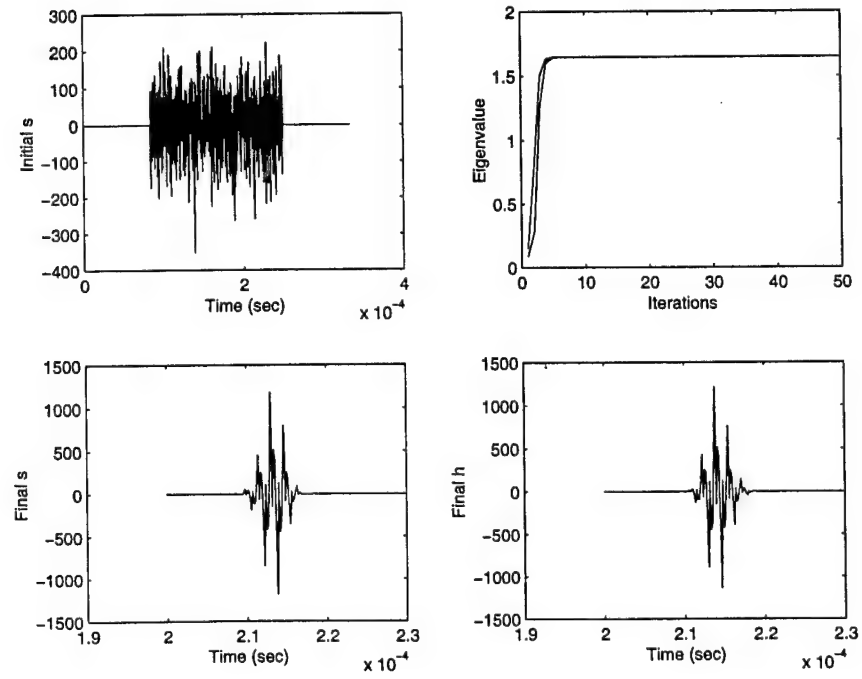
The results in Figs. 7.13 and 7.14 are obtained with $f = 6250$Hz and $\beta = 0.3$ for single-path fading. The spread of the signal pulses increases with the synchronization uncertainty. At the same time, the largest eigenvalue or the SNR decreases as the signal spread increases.

7.4.2 No Fading

It is interesting to see how the synchronization delay affects the CP. The results can be obtained by setting $\beta = 0$ and the coefficients of the channel autocorrelation matrix, $\mathbf{r}$, to some constant values, which is equivalent to having a constant autocorrelation of the channel. Fig. 7.15(a) depicts the significance of the Match Filter Theory. That is, any

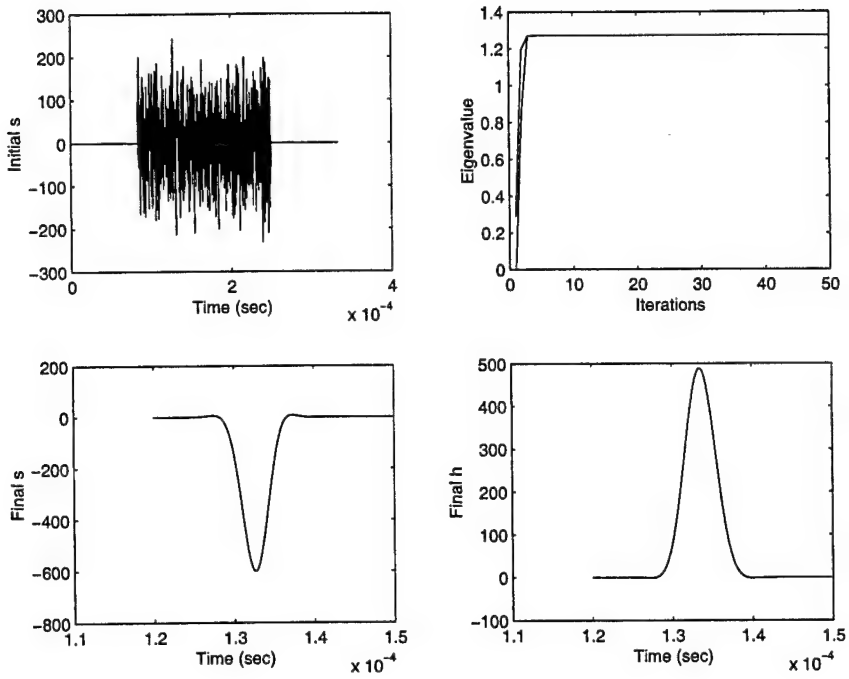


(a) Perfect synchronization, single path.

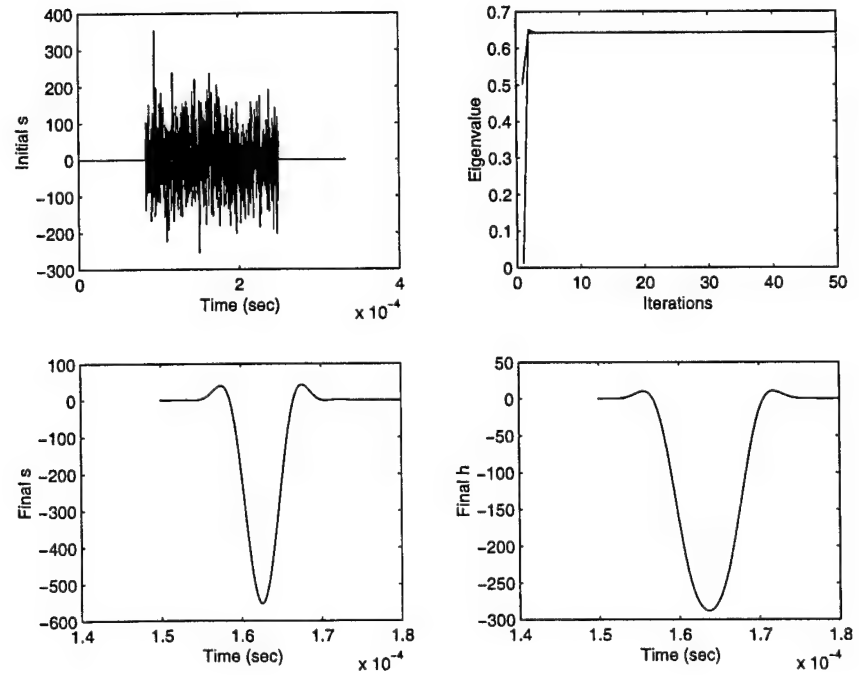


(b) TU, perfect synchronization

Fig. 7.11: Effect of multiple paths I.

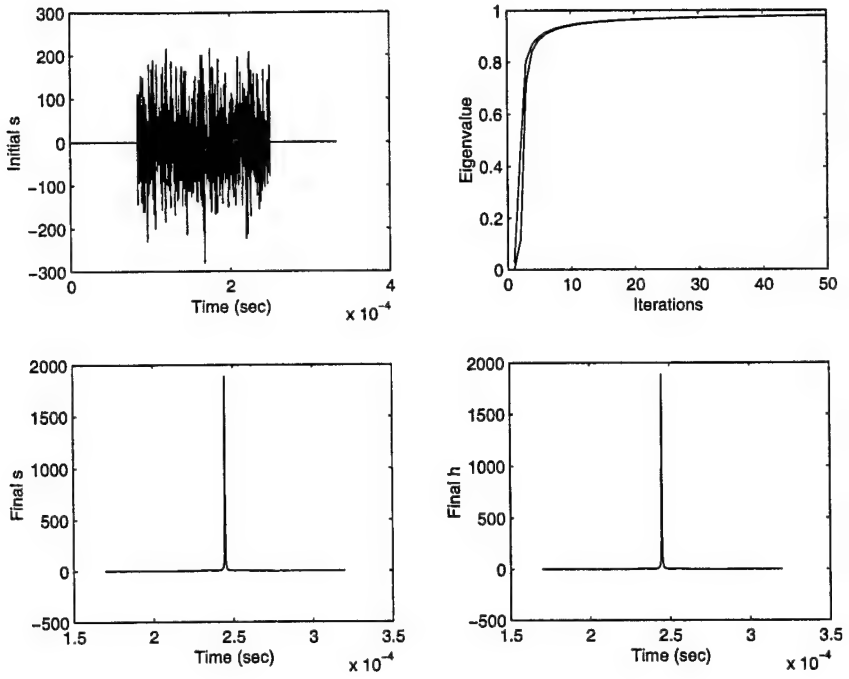


(a) TU, Gaussian-distributed delay with $\sigma = 1.124 \times 10^{-6}$sec

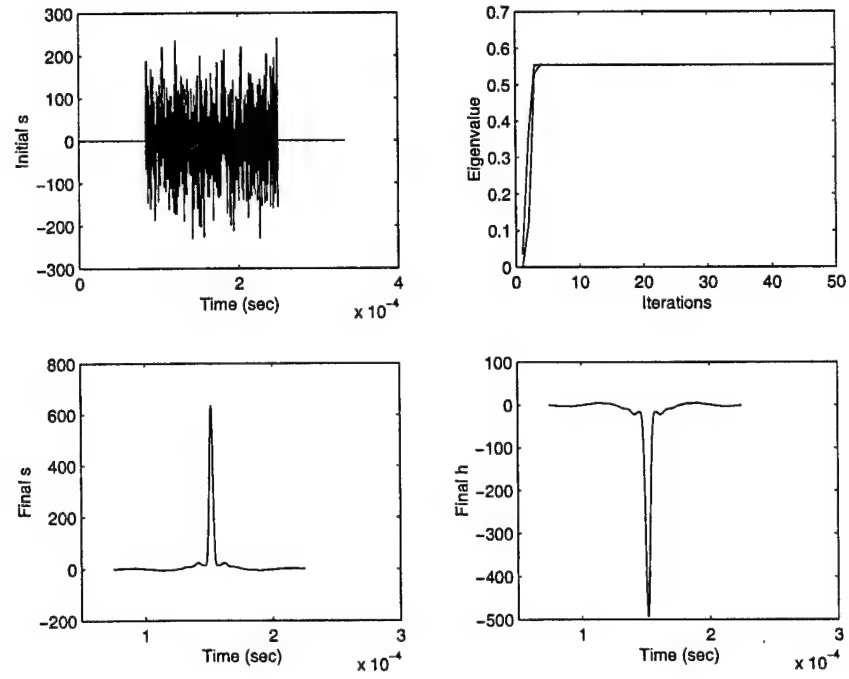


(b) TU, uniform distribution with $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec

Fig. 7.12: Effect of multiple paths II.

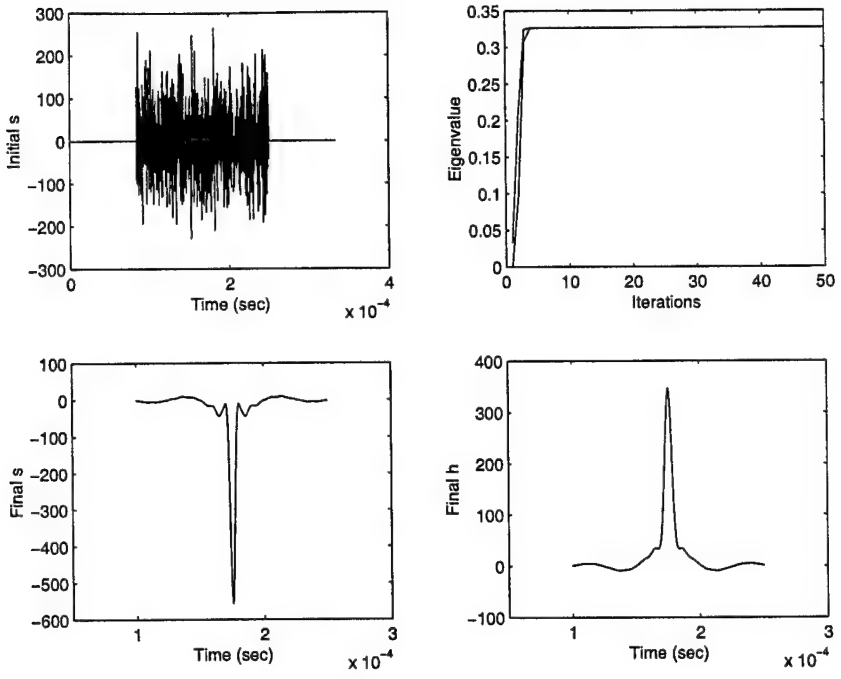


(a) Perfect synchronization

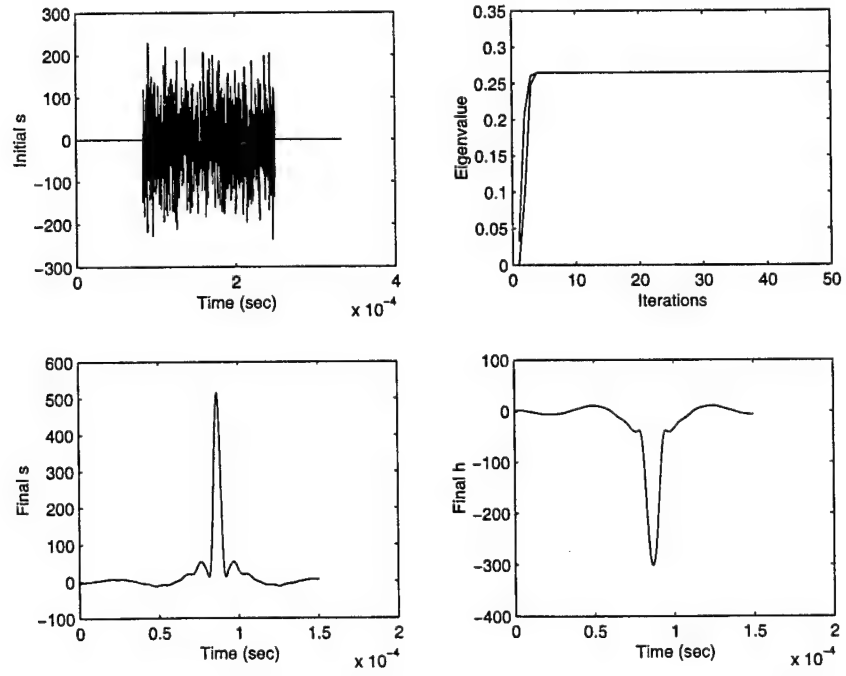


(b) Gaussian-distributed delay with $\sigma = 1.124 \times 10^{-6}$sec

Fig. 7.13: Effect of synchronization uncertainty I.



(a) Gaussian-distributed delay with $\sigma = 2.225 \times 10^{-6}$sec



(b) Uniformly distributed with $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec

Fig. 7.14: Effect of synchronization uncertainty II.

signal-filter pair will perform equally well for a nonfading-single-path channel, as long as the signal-filter pair is matched. On the other hand, Fig. 7.15(b) shows that if there is a synchronization uncertainty, the CP is shaped by the distribution of delay. The effect of synchronization uncertainty on the shaping of CPs also occurs in multipath channels as shown in Figs. 7.16(b) and 7.17(b). Furthermore, for the cases of multipath channels, a CP is also controlled by the path delays, even where there is no synchronization uncertainty as shown in Figs. 7.16(a) and 7.17(a).

7.5 Signal Sets

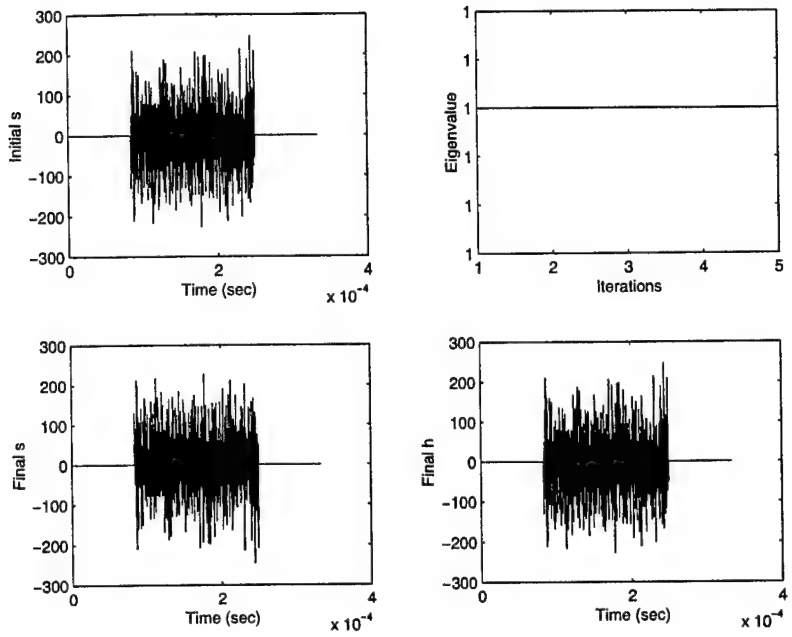
We discussed in section 4.4 that there are two alternatives to apply the results from this dissertation to obtain signal sets for digital communication. These alternatives are: First, use the CPs generated with Algorithm 4 in a pulse-position modulation. Second, use the CCPs generated with Algorithm 5. In this section, we compare the performance of the signal/filter pairs obtained by using our approach with some traditional waveforms, i.e., the flat-top pulses and the raised-cosine pulses.

7.5.1 CCPs

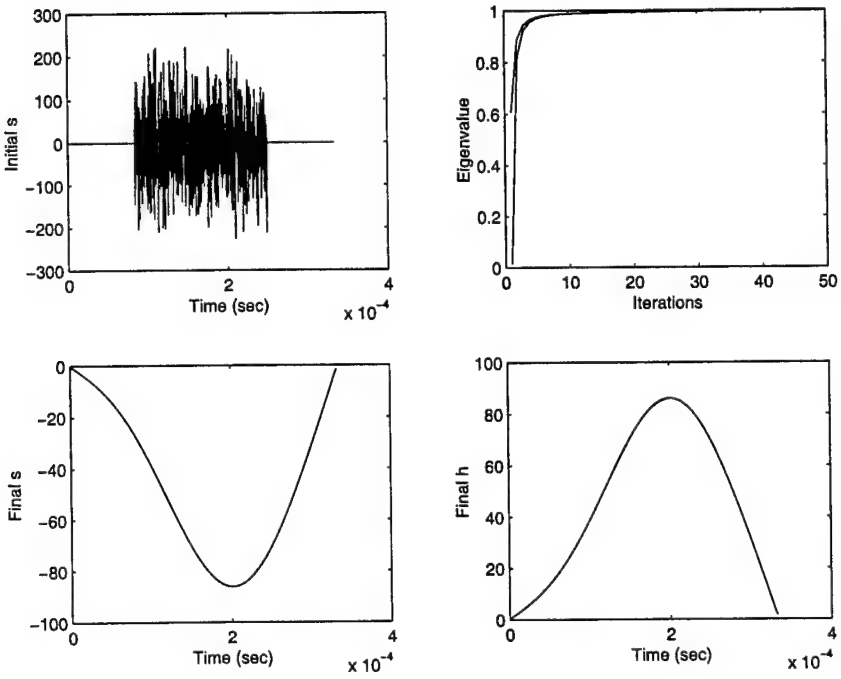
Because of the slow convergence of Algorithm 5, all plots in this section are obtained with 512-point-sample vectors. We use $\delta T = 1.626 \times 10^{-7} s$. This is equivalent to a data rate of 12,012 bits per second. In all of the plots, we assume single path fading channel with perfect synchronization.

Figs. 7.18 and 7.19 show the change in signal/filter of CCPs according to different values of β. In each of these figures, the group of four pictures of CCPs is arranged into two rows. On the top row, the picture on the left-hand side shows $\mathbf{s}_0$ while the one on the right-hand side shows $\mathbf{h}_0$. On the bottom row, the picture on the left-hand side shows $\mathbf{s}_1$ while the one on the right-hand side shows $\mathbf{h}_1$. We can observe from all these plots that the signal and filter functions are getting smoother as β increases.

Fig. 7.20 shows the performance plots of probability-of-error versus SNR. On the performance plots of joint CPs, we also show that of binary flat-top pulse and raised-

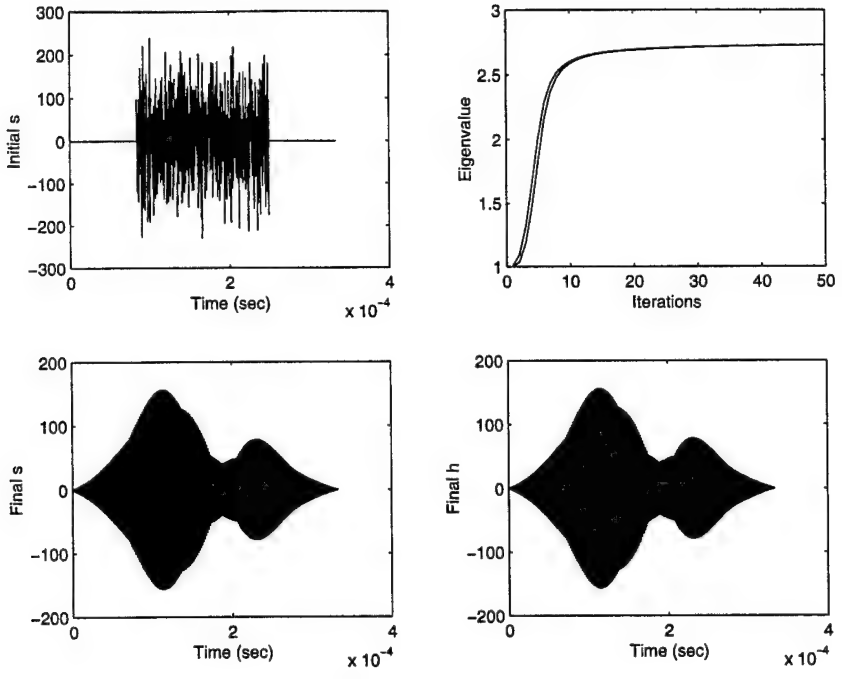


(a) Single path, perfect synchronization

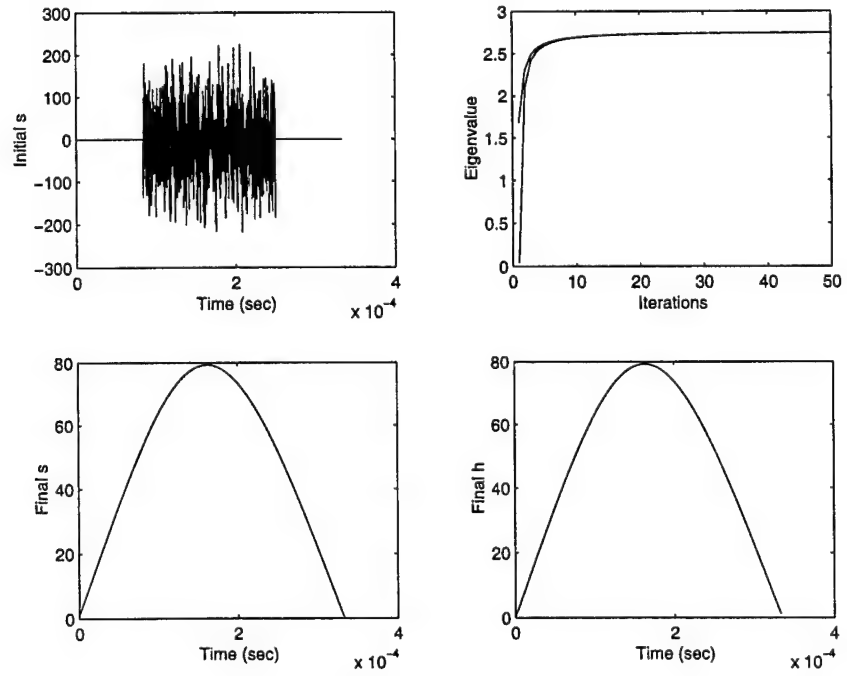


(b) Single path, uniformly distributed delay with $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec

Fig. 7.15: CP for channels without fading I.

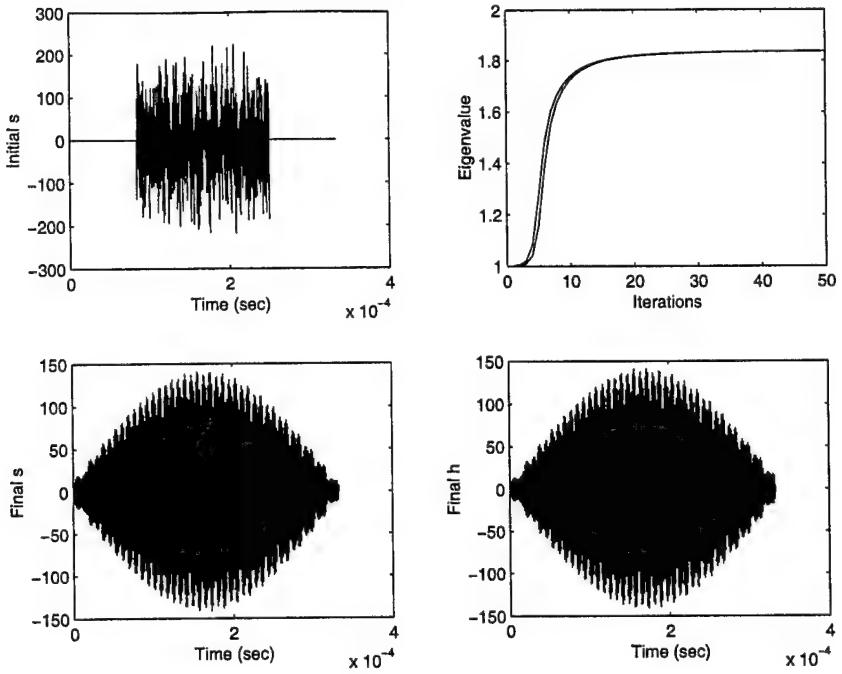


(a) TU, perfect synchronization

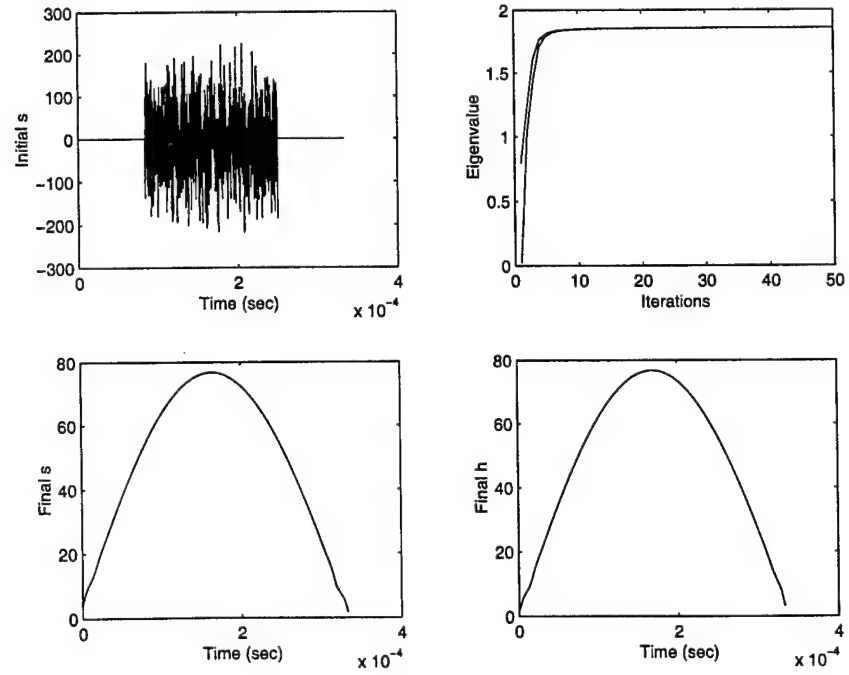


(b) TU, uniformly distributed delay with $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec

Fig. 7.16: CP for channels without fading II.

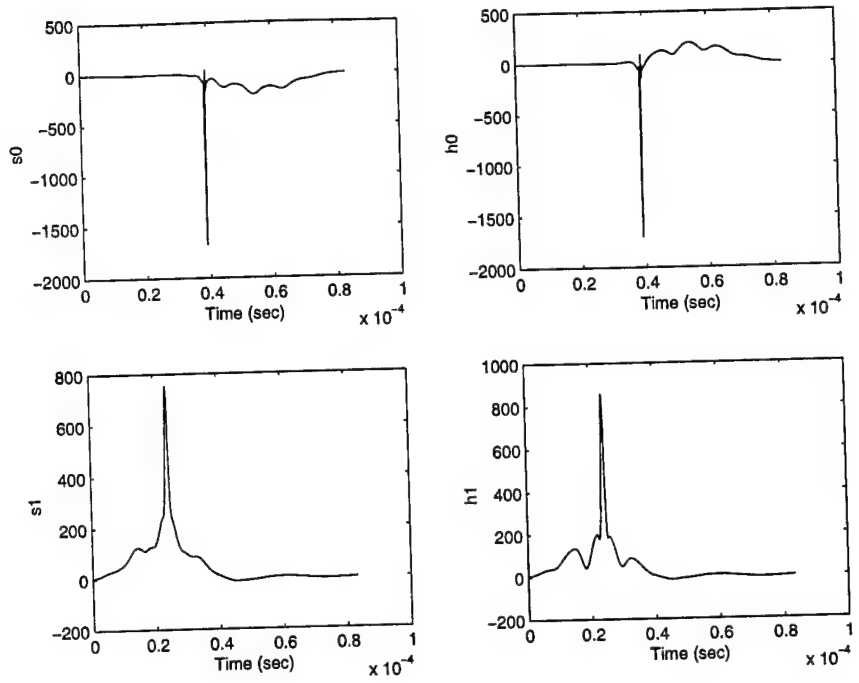


(a) HT, perfect synchronization

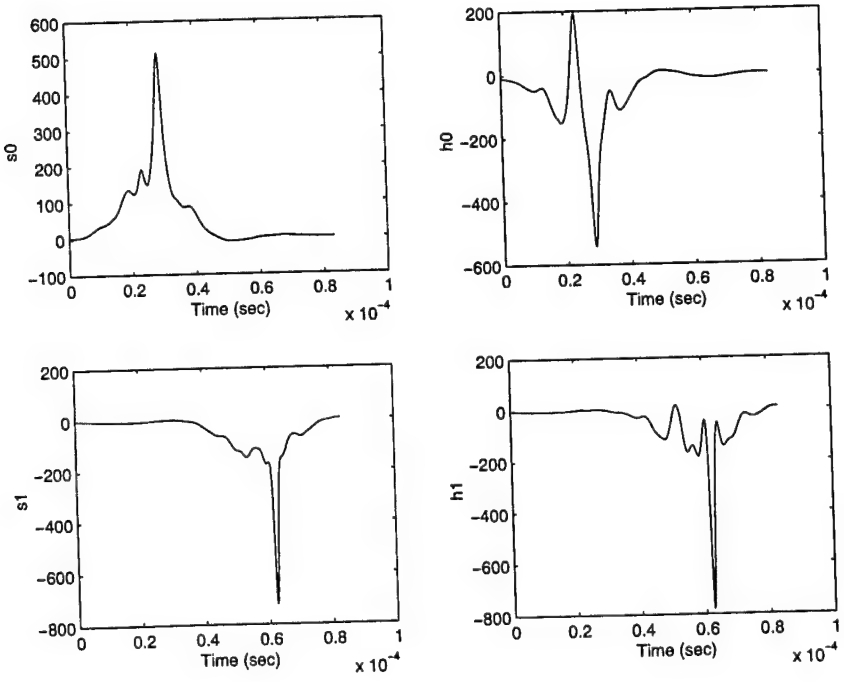


(b) HT, uniformly distributed delay with $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec

Fig. 7.17: CP for channels without fading III.

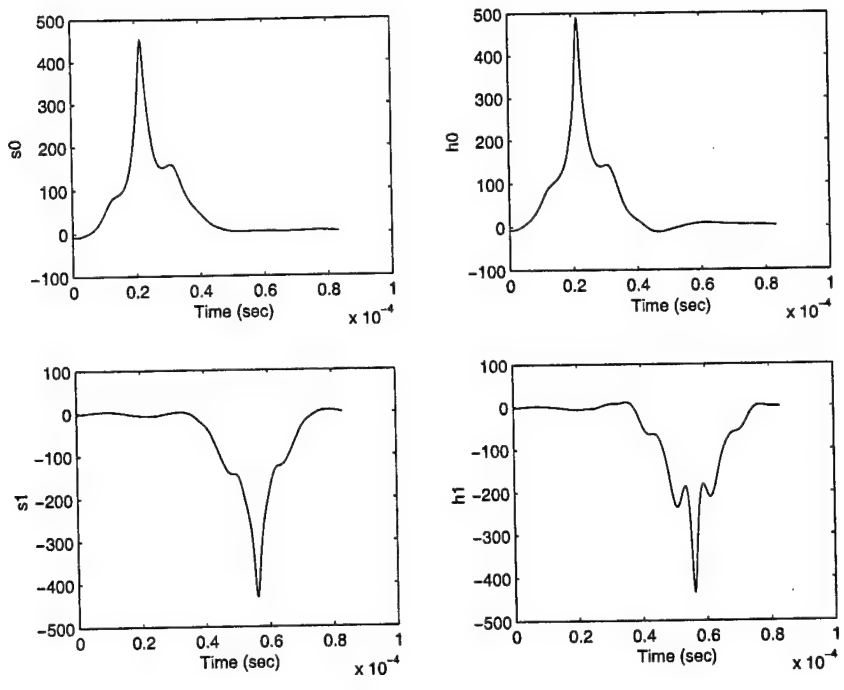


(a) $\beta = 0.1$

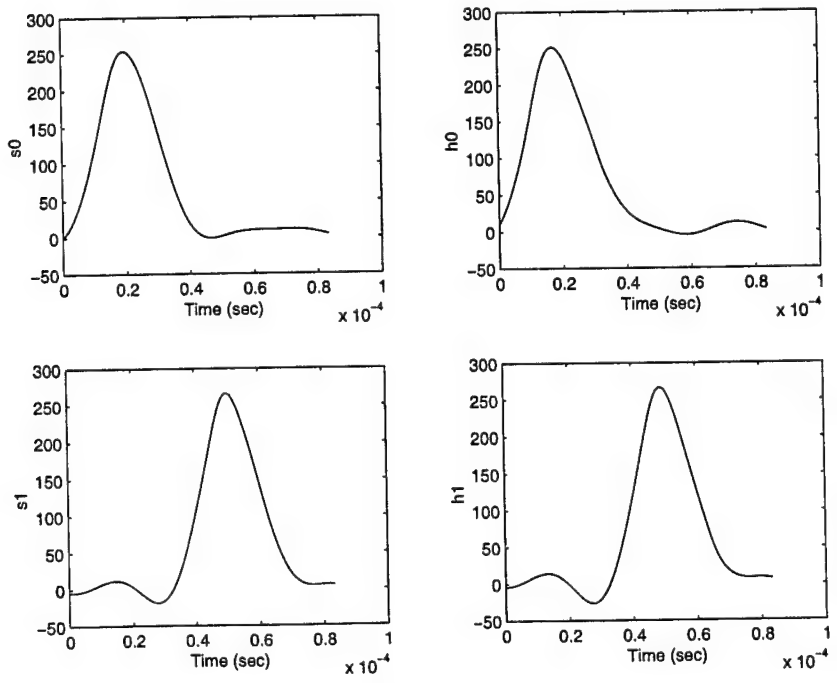


(b) $\beta = 0.3$

Fig. 7.18: CCPs for single path, perfect synchronization $\beta = 0.1$ and $\beta = 0.3$.



(a) $\beta = 0.5$



(b) $\beta = 0.7$

Fig. 7.19: CCPs for single path, perfect synchronization $\beta = 0.5$ and $\beta = 0.7$.

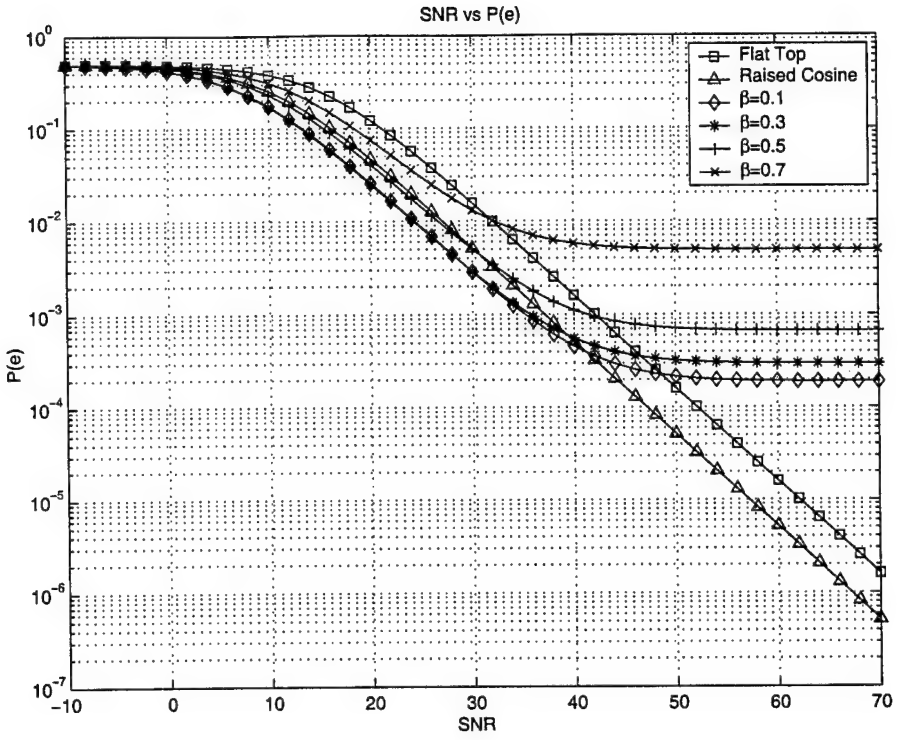


Fig. 7.20: SNR vs. $P(e)$, CCPs for single path, perfect synchronization.

cosine pulse in pulse position modulation (PPM) pairs with a support of $8.35 \times 10^{-5}s$ in the same channel for comparison. The observations are: First, for CCPs, the smaller the β, the better the performance. Second, there are error floors for the CCPs. Third, before reaching their error floor, CCPs with $\beta = 0.1$ and $\beta = 0.3$ have a performance gain of 2 dB to 3 dB over the raised-cosine pulse, and have a 8 dB to 9 dB gain over the flat-top pulse. CCPs with $\beta = 0.5$ have a performance marginally better than that of the raised-cosine pulse, while that with $\beta = 0.7$ have a performance loss of around 2 dB over that of the raised-cosine pulse.

From all of the experiments we did for different fading channels, there is a similar trend of relationships between the performance of CCPs, raised-cosine pulse and flat-top pulse as observed above. Another observation is different initial vectors can generate CCPs with very different performance. We used four different initial vectors for each set of channel parameters and ran Algorithm 5 for 60 iterations with each of these initial vectors. The best CCPs obtained with each of the different initial vector can differ by more than ten

folds in terms of probability-of-error for the same set of channel parameters. Considering the small number of experiments we did, the possible best CCPs may perform far better than these CCPs we found.

7.5.2 PPM with CPs

Results presented in Figs. 7.21 through 7.24 are obtained with $f = 12,500$Hz and a data rate of 1500 bits per second. The length of the sample vectors are 4096, with $\delta T = 1.626 \times 10^{-7} s$. From all the plots, we observe: first, a 5 to 10 dB gain in performance for the raised-cosine pulse over the flat-top pulse for different fading and synchronization situations; and second, CPs with $\beta = 0.1$ through $\beta = 0.5$ have very similar performance while the one with $beta = 0.7$ is generally inferior to the other CPs. Figure 7.21 shows the performance of the CPs for TU multipath fading model with a uniformly distributed synchronization delay such that $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec. The waveforms of these CPs are shown in Figs. 7.7 and 7.8. There is an 8 to 9 dB gain in performance of all these CPs over the raised-cosine pulse.

Figure 7.22 shows the performance of the CPs for HT multipath fading model without synchronization delay whose waveforms are shown in Figs. 7.7 and 7.8. We also observe a 10 to 12 dB gain in performance of the CPs over the raised-cosine pulse. Now, there is less than 2 dB difference between the CP with $\beta = 0.7$ and the other CPs.

Figure 7.23 shows the performance of the CPs for single path fading with a Gaussian-distributed synchronization delay ($\sigma = 1.124$), whose waveforms for different β's are shown in Figs. 7.5 and 7.6. In this figure, we observe that the flat-top pulse reaches its error floor at the SNR of 65 dB. We also observe a 10 to 13 dB gain in performance of the CPs over the raised-cosine pulse. This time, there is a 3 dB difference between the CP with $\beta = 0.7$ and the other CPs.

Figure 7.24 shows the performance of the CPs for single path fading with a perfect synchronization whose waveforms for different values of β are shown in Figs. 7.3 and 7.4. We observe a 10 to 15 dB gain in performance of the CPs over the raised-cosine pulse, while the CP with $\beta = 0.7$ has a 5 dB difference from the other CPs.

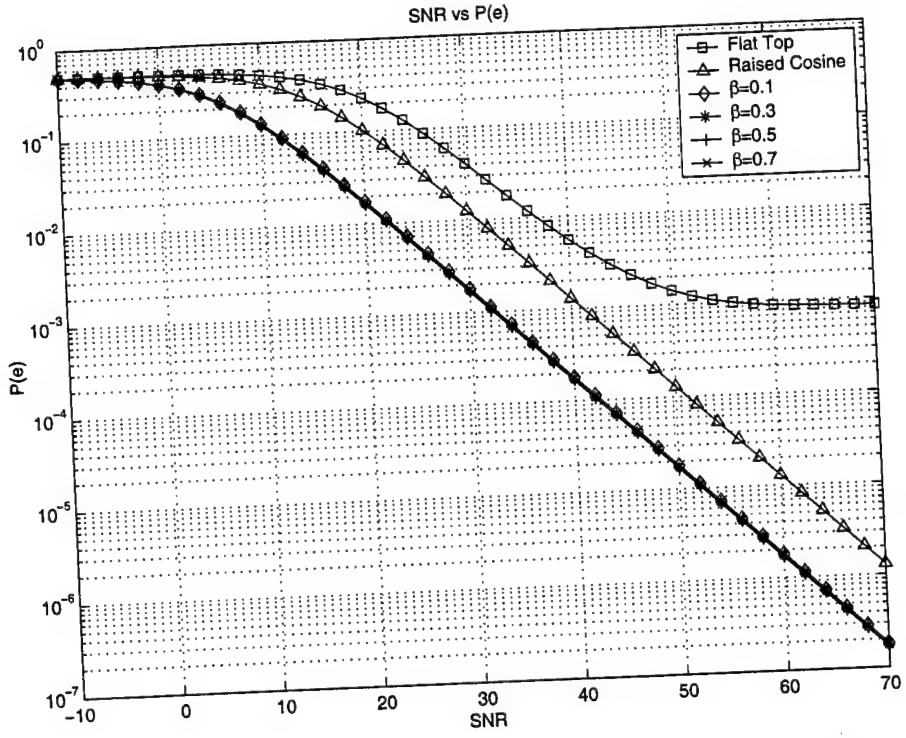


Fig. 7.21: TU, uniformly distributed, $\mathcal{U}(-4.15 \times 10^{-6}, 4.15 \times 10^{-6})$sec.

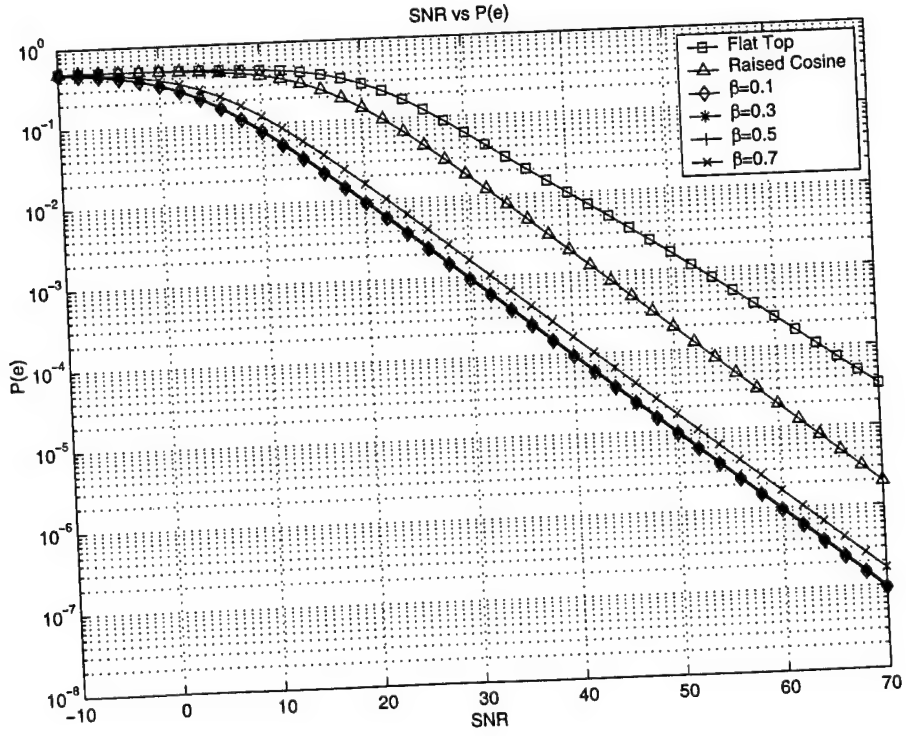


Fig. 7.22: HT, perfect synchronization.

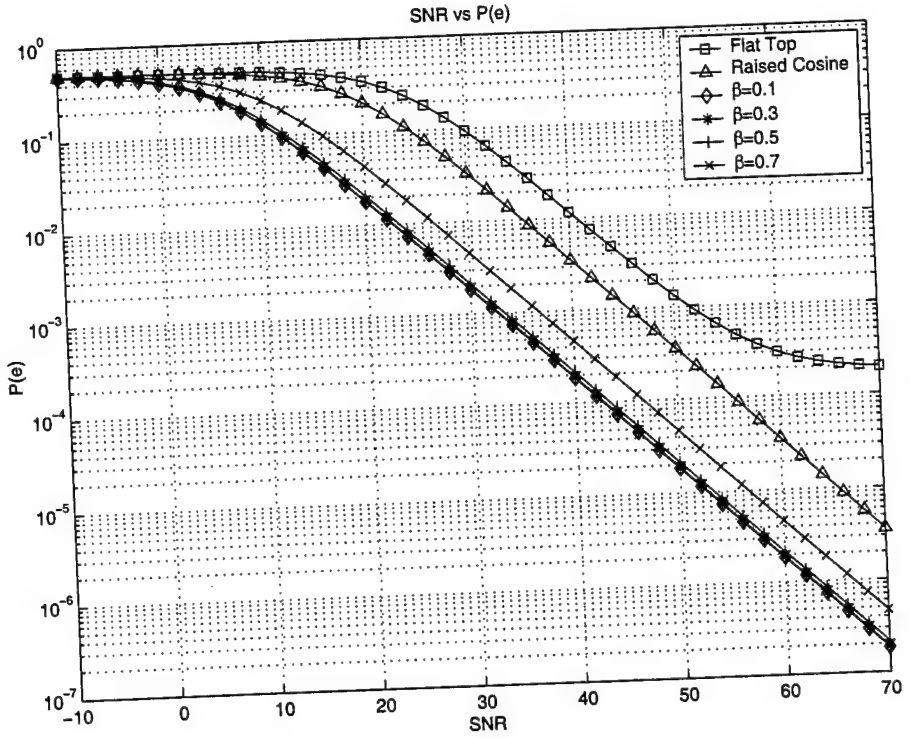


Fig. 7.23: Single path, Gaussian-distributed with $\sigma = 1.124$.

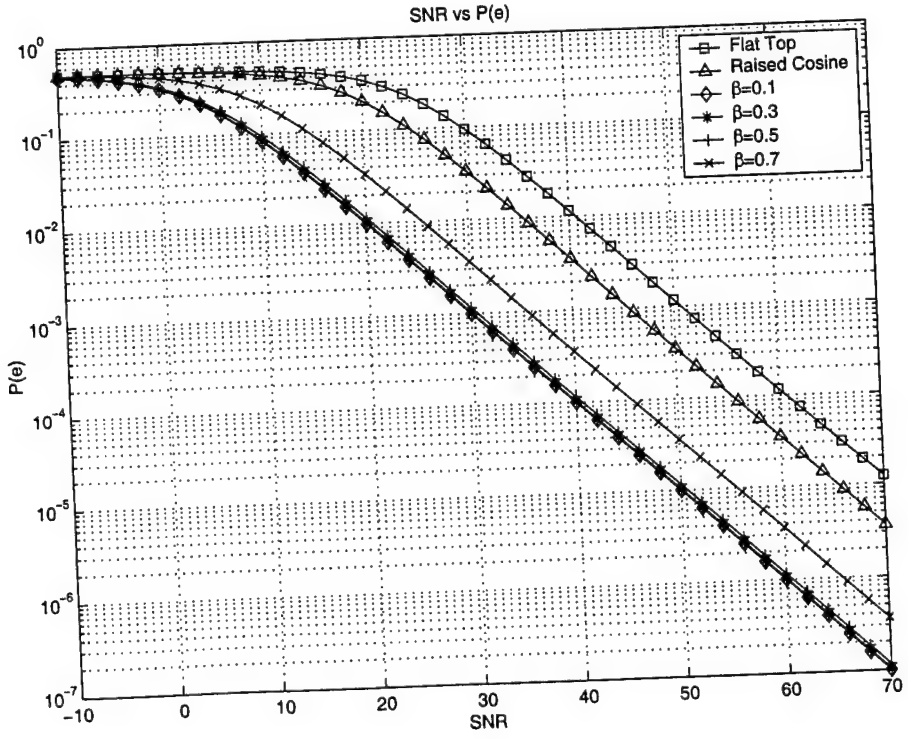


Fig. 7.24: Single path, perfect synchronization.

7.6 The Convergence of Algorithms

In some cases, Algorithms 4 and 5 do not converge. Among all unknown causes of the failure for these algorithms to converge, there are two major possible causes. Firstly, the power method or the conjugate gradient method may not have converged rapid enough to the correct eigenvalue/eigenvector by the time the preset number of iterations has been exhausted. This situation may occur if the largest eigenvalues are too close together for the power method, or the smallest eigenvalues are too close together for the conjugate gradient method (see discussion in section 6.4). Secondly, the averaging effect of synchronization uncertainty counteracts with the convergence.

Chapter 8
Conclusion and Discussion

In this chapter, the contributions made in this dissertation are summarized. Directions for future study are discussed.

8.1 Summary

In this communication theory research, we were able to characterize the optimal detector and signal by maximizing the SNR for communication through multiplicative channels, which may be slow or fast fading. In addition, this characterization is general enough to include single-path and multipath fading with different synchronization-delay profiles.

The design criterion of maximizing SNR with fixed energy leads to a constrain optimization problem. The optimization was performed by using the Lagrange multipliers in a function space. We were able to characterize the optimal detector as the eigenfunction corresponding to the largest eigenvalue of an operator with kernel being the product of the channel autocorrelation function and a transformation of the transmitted signal. The transformation of the transmitted signal depends on the multipath and the synchronization-delay profiles.

We then derived the optimal signal and detector as a pair, this time optimizing the two variables simultaneously using the Lagrange multipliers. The conclusion we obtained was that the signal and the filter must be similar to each other in amplitude, but they might differ in sign. However, the resulting expression did not provide a way to determine the optimal functions physically besides a list of properties as in section 3.4.

Based on the symmetric role of the filter and signal in a communication system, we derived an iterative algorithm (Algorithm 1) which allowed us to generate an optimal consonant pair from random seeds. This iterative algorithm contains two alternating steps:

1. an optimal detector to generate an updated *filter* function for the previously generated

signaling function; and 2. an optimal detector to generate an updated *signaling* function for the previously generated *filter* function.

For single-path fading, the above iterative algorithm always generates consonant pairs that approach delta distributions. Physically this is very reasonable. If there is an uncertainty in the channel status, a transmitted signal should use as little time of the channel as possible. However, a function that is close to the delta distribution is not practical, because it requires infinite band width to transmit. So, we introduce a frequency containment constraint to Algorithm 1 and derived Algorithm 2.

In order to communicate over a fast fading channel, the use of amplitude and phase modulation is inappropriate. Something akin to orthogonal signaling is necessary. A procedure, Algorithm 3 was developed to generate two CCPs such that the corresponding functions were nearly orthogonal among the conjugate pairs. Algorithm 3 required two copies of Algorithm 2 being executed in parallel with the additional requirement that the updated consonant pair from one execution was related to the eigenfunction corresponding to the smallest eigenvalue of the other execution.

Furthermore, we derived Algorithms 4 and 5, the discrete version of Algorithms 2 and 3, on which we did all the experiments.

To help us compare CCPs generated by our procedure with other standard signaling waveforms, we also derived the density and distribution functions for quadratic forms of complex normal vectors. The equation we obtained can be evaluated by using the eigenvalues of a single matrix. This matrix is the product of the correlation matrix of the normal vector and the Hermitian of the quadratic form. This is an extension of previous results as most of these results would require us to evaluate some infinite sequence or only applicable to some specific values. A partition of unity found in the evaluation of the distribution function at point zero is useful in simplifying some expressions used in standard digital communication reference books.

8.2 Future Work

This dissertation provides a new approach to signal and detector design for communication through fading channels. There are many directions that can be taken from this point both theoretically and practically. Here we list a few of them.

1. Convergence of the algorithm: It is easy to show the algorithm for single-path fading channels when there is no synchronization uncertainty with $\beta = 0$ or $\beta = 1$ has a monotonic increasing largest eigenvalue. The convergence of the algorithm for the above and other situations remains unknown. Knowing the limit of the convergence of the algorithm will help to extend the analysis by applying similar technique to more the complex models.

2. Quadratic forms: The development of the expression of the density and the distribution functions of quadratic forms with zero mean normal vectors is sufficient for this dissertation. However, to make the expression for the density and the distribution functions truly useful, e.g., applicable to the slow fading situations, an extension of these equations to include nonzero mean normal vectors is necessary and is highly valued.

3. Error floor: The phenomenon of error floor is discussed in the literature, including those written by Hansson [108] and Simon [92]. The concept of error floor is referred to as "irreducible bit error probability" or "error floor." The error floor is the limiting behavior of $P(e) > 0$ as the SNR approaches infinity for some signaling schemes with their corresponding receivers. Simon pointed out that for M-ary Differential phase-shift-keying, the error floor exists in any channel with fading autocorrelation that is not equal to one (fast fading). Hansson, on the other hand, showed that for orthogonal signal/filter pairs, the error floor does not exists. Results from section 5.5 can be used to extend the characterization of the signal/filter pairs that do not have an error floor.

4. Consonant pairs for slow/nonfading channels: We observed from the experiments that there are relationships between the synchronization uncertainty and the shapes of a consonant pair. This is an area that may have significant practical importance, as synchronization uncertainty is unavoidable in communication.

5. Analysis of the consonant pair: We can observe that consonant pairs of different channels share many similar properties. If we can analyze and successfully represent the consonant pairs of most common channels with some specific family of functions, e.g., wavelets, we have a flexible communication system that can easily adapt to these common channels.

References

[1] R. W. Lucky, *Silicon Dreams: Information, Man, and Machine.* New York: St. Martin's Press, 1989.

[2] K. Sato and S. Okamoto, "Photonic transport technologies to create robust backbone networks," *IEEE Comm. Mag.*, vol. 37, pp. 78–87, Aug. 1999.

[3] M. Zeng, A. Annamalai, and V. K. Bhargava, "Recent advances in cellular wireless communications," *IEEE Comm. Mag.*, vol. 37, pp. 128–138, Sept. 1999.

[4] M. Budagavi, W. R. Heinzelman, J. Webb, and R. Talluri, "Wireless MPEG-4 video communication on DSP chips," *IEEE Sig. Proc. Mag.*, vol. 17, pp. 36–53, Jan. 2000.

[5] M. Wittig, "Large-capacity multimedia satellite systems," *IEEE Comm. Mag.*, vol. 35, pp. 44–49, July 1997.

[6] A. Dutta-Roy, "Fixed wireless routes for internet access," *IEEE Spectrum*, vol. 36, pp. 61–69, Sept. 1999.

[7] C. O'Malley, "What's next: Computing - it's an internet world," *Popular Science*, pp. 46–61, Mar. 2000.

[8] T. S. Rappaport, B. D. Woerner, and J. H. Reed, eds., *Wireless Personal Communications: The Evolution of Personal Communications Systems.* Boston, MA: Kluwer Academic Publishers, 1996.

[9] W. C. Jakes, ed., *Microwave Mobile Communications.* New York: Wiley, 1974.

[10] A. F. Naguib, N. Seshadri, and A. R. Calderbak, "Increasing data rate over wireless channels," *IEEE Sig. Proc. Mag.*, vol. 17, pp. 76–92, May 2000.

[11] R. C. V. Macario, *Cellular Radio: Principles and Design.* New York: McGraw-Hill, 2 ed., 1997.

[12] B. Sklar, "Rayleigh fading channels in mobile digital communication systems, part II: Mitigation," *IEEE Commun. Mag.*, vol. 35, pp. 148–155, Sept. 1997.

[13] K. Brayer, "HF data transmission: Lessons from the past, directions for the future," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, pp. 90–101, Feb. 1987.

[14] R. Price, "The detection of signals perturbed by scatter and noise," *IRE Trans. Inform. Theory*, vol. PGIT-4, pp. 163–170, Sept. 1954.

[15] R. Price, "Optimum detection of random signals in noise, with application to scatter-multipath communication," *IRE Trans. Inform. Theory*, vol. IT-2, pp. 125–135, Dec. 1956.

[16] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering.* New York: John Wiley and Sons, Inc., 1967.

[17] P. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Comm. Sys.*, vol. CS-11, pp. 360–393, Dec. 1963.

[18] R. S. Kennedy, *Fading Dispersive Communication Channels.* New York: Wiley-Interscience, 1969.

[19] K. Brayer, ed., *Data Communications via Fading Channels.* New York: IEEE Press, 1975.

[20] S. Stein, "Fading channel issues in system engineering," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, pp. 68–89, Feb. 1987.

[21] J. G. Proakis, *Digital Communications.* New York: McGraw Hill, 3 ed., 1995.

[22] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[23] B. Sklar, "Rayleigh fading channels in mobile digital communication systems, part I: Characterization," *IEEE Commun. Mag.*, vol. 35, pp. 136–146, Sept. 1997.

[24] V. Poor and G. W. Wornell, eds., *Wireless Communications.* Upper Saddle River, NJ: Prentice Hall, 1998.

[25] "Scatter issue," *Proceedings of the IEEE*, vol. 43, Dec. 1955.

[26] A. B. Glenn and G. Lieberman, "Performance of digital communications systems in an arbitrary fading rate and jamming environments," *IEEE Trans. Comm. Sys.*, vol. CS-11, pp. 57–68, Mar. 1963.

[27] P. Bello, "A troposcatter channel model," *IEEE Trans. Comm. Tech.*, vol. 17, pp. 130–137, Apr. 1969.

[28] "West ford belt issue," *Proceedings of the IEEE*, vol. 52, May 1964.

[29] G. F. Montgomery, "Message error in diversity frequency-shift reception," *Proceedings of the IRE*, vol. 42, pp. 1184–1187, July 1954.

[30] J. Pierce, "Theoretical diversity improvement in frequency-shift keying," *Proceedings of the IRE*, vol. 46, pp. 903–910, May 1958.

[31] D. Brennan, "Linear diversity combining techniques," *Proceedings of the IRE*, vol. 47, pp. 1075–1102, June 1959.

[32] J. Pierce, "Theoretical limitations of frequency and time diversity for fading binary transmissions," *IRE Trans. Commun. Sys.*, vol. CS-9, pp. 186–189, June 1961.

[33] G. L. Turin, "On optimal diversity reception," *IRE Trans. Inform. Theory*, vol. IT-7, pp. 154–166, July 1961.

[34] M. V. Clark, L. J. Greenstein, W. K. Kennedy, and M. Shafi, "Matched filter performance bounds for diversity combining receivers in digital mobile radio," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 356–362, Nov. 1992.

[35] D. Bouras, P. Mathiopoulos, and D. Makrakis, "Optimal detection of coded differentially encoded QAM and PSK signals with diversity reception on correlated fast Ricean fading channels," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 245–258, Aug. 1993.

[36] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation - part I: Coherent detection," *IEEE Trans. Comm.*, vol. 45, pp. 563–572, May 1997.

[37] A. J. Viterbi and I. M. Jacobs, "Advances in coding and modulation for noncoherent channels affected by fading, partial band, and multiple-access interference," in *Advances in Communication Systems* (A. J. Viterbi, ed.), New York: Academic, 1975.

[38] D. Chase, "Digital signal design concepts for a time-varying Ricean channel," *IEEE Trans. Comm.*, vol. COM-24, pp. 164–172, Feb. 1976.

[39] J. N. Pieper, J. G. Proakis, R. R. Reed, and J. K. Wolf, "Design of efficient coding and modulation for a Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 457–468, July 1978.

[40] I. Rahman, *Bandwidth Constrained Signal Design for Digital Communication over Rayleigh Fading Channels and Partial Band Interference Channels.* PhD thesis, Northeastern University, Boston, Mass., 1981.

[41] G. W. Wornell, "Spread-response precoding for communication over fading channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 488–501, Mar. 1996.

[42] P. Y. Kam, "Bit-error probabilities of 2 and 4 DPSK with nonselective Rayleigh fading, diversity reception, and correlated Gaussian interference," *IEEE Trans. Comm.*, vol. 45, pp. 400–403, Apr. 1997.

[43] E. A. Neasmith and N. C. Beauieu, "New results on selection diversity," *IEEE Trans. Comm.*, vol. 46, pp. 695–704, May 1998.

[44] J. Lu, T. T. Tjhung, and C. C. Chai, "Error probability performance of l-branch diversity reception of MQAM in Rayleigh fading," *IEEE Trans. Comm.*, vol. 46, pp. 179–181, Feb. 1998.

[45] M. K. Simon and M.-S. Alouini, "A unified performance analysis of digital communication with dual selective combining diversity over correlated Rayleigh and Nakagami-m fading channels," *IEEE Trans. Comm.*, vol. 47, pp. 33–43, Jan. 1999.

[46] D. Rainish, "Diversity transform for fading channels," *IEEE Trans. Comm.*, vol. 44, pp. 1653–1661, Dec. 1996.

[47] A. M. Sayeed and B. Aazhang, "Joint multipath-doppler diversity in mobile wireless communications," *IEEE Trans. Comm.*, vol. 47, pp. 123–132, Jan. 1999.

[48] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation - part III: Co-channel interference," *IEEE Trans. Comm.*, vol. 45, pp. 809–818, July 1997.

[49] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[50] V. M. DaSilva and E. S. Sousa, "Fading-resistant modulation using several transmitter antennas," *IEEE Trans. Comm.*, vol. 45, pp. 1236–1244, Oct. 1997.

[51] A. Shah and A. M. Haimovich, "Performance analysis of optimum combining in wireless communications with Rayleigh fading and cochannel interference," *IEEE Trans. Comm.*, vol. 46, pp. 473–479, Apr. 1998.

[52] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Comm.*, vol. 47, pp. 1458–1461, Oct. 1999.

[53] G. M. Vitetta, U. Mengali, and D. P. Taylor, "Double-filter differential detection of PSK signals transmitted over linearly time-selective Rayleigh fading channels," *IEEE Trans. Comm.*, vol. 47, pp. 239–247, Feb. 1999.

[54] M. P. C. Fossorier and S. Lin, "Soft decision decoding of linear block codes based on ordered statistics for the Rayleigh fading channel with coherent detection," *IEEE Trans. Comm.*, vol. 45, pp. 12–14, Jan. 1997.

[55] M. Visintin, "Differential PSK block demodulation over a flat correlated Rayleigh-fading channel," *IEEE Trans. Comm.*, vol. 45, pp. 9–11, Jan. 1997.

[56] S. Buzzi, E. Conte, and M. Lops, "Optimum detection over Rayleigh-fading, dispersive channels, with non-Gaussian noise," *IEEE Trans. Comm.*, vol. 45, pp. 1061–1069, Sept. 1997.

[57] F. Patenaude, J. H. Lodge, and Y.-Y. Chouinard, "Noncoherent diversity reception over Nakagami-fading channels," *IEEE Trans. Comm.*, vol. 46, pp. 985–991, Aug. 1998.

[58] L. M. Davis, I. B. Colings, and R. J. Evans, "Coupled estimators for equalization of fast-fading mobile channels," *IEEE Trans. Comm.*, vol. 46, pp. 1262–1265, Oct. 1998.

[59] H.-N. Lee and G. J. Pottie, "Fast adaptive equalization/diversity combining for time-varying dispersive channels," *IEEE Trans. Comm.*, vol. 46, pp. 1146–1162, Sept. 1998.

[60] S. D. Fina and G. E. Corraza, "Bayesian approach for erasure insertion in frequency-hop multiple-access communications with selective fading," *IEEE Trans. Comm.*, vol. 48, pp. 282–289, Feb. 2000.

[61] R. Cusani and J. Mattila, "Equalization of digital radio channels with large multipath delay for cellular land mobile applications," *IEEE Trans. Comm.*, vol. 47, pp. 348–351, Mar. 1999.

[62] D. K. Borah and B. D. Hart, "A robust receiver structure for time-varying, frequency-flat, Rayleigh fading channels," *IEEE Trans. Comm.*, vol. 47, pp. 360–364, Mar. 1999.

[63] W. S. Leon, U. Mengali, and D. P. Taylor, "Equalization of linearly frequency-selective fading channels," *IEEE Trans. Comm.*, vol. 45, pp. 1501–1503, Dec. 1997.

[64] W. S. Leon and D. P. Taylor, "An adaptive receiver for the time- and frequency-selective fading channel," *IEEE Trans. Comm.*, vol. 45, pp. 1548–1555, Dec. 1997.

[65] Y. A. Chau and J.-K. Wang, "Spectral-estimation-based acquisition for frequency-hopping spread-spectrum communications in a nonfading or Rayleigh-fading channel," *IEEE Trans. Comm.*, vol. 45, pp. 445–455, Apr. 1997.

[66] M. L. McCloud and L. L. Scharf, "Interference estimation with applications to blind multiple-access communication over fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 947–961, May 2000.

[67] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Trans. Comm.*, vol. 16, pp. 1423–1436, Oct. 1998.

[68] S. Bhashyam, A. M. Sayeed, and B. Aazhang, "Time-selective signaling and reception for communication over multipath fading channels," *IEEE Trans. Comm.*, vol. 48, pp. 83–94, Jan. 2000.

[69] A. J. Goldsmith and S.-G. Chua, "Adaptive coded modulation for fading channels," *IEEE Trans. Comm.*, vol. 46, pp. 595–602, May 1998.

[70] T. Muller and H. Rohling, "Channel coding for narrow-band Rayleigh fading with robustness against changes in doppler spread," *IEEE Trans. Comm.*, vol. 45, pp. 148–151, Feb. 1997.

[71] F. Babich and G. Lombardi, "Statistical analysis and characterization of the indoor propagation channel," *IEEE Trans. Comm.*, vol. 48, pp. 455–464, Mar. 2000.

[72] A. J. Coulson, A. G. Williamson, and R. G. Vaughan, "A statistical basis for lognormal shadowing effects in multipath fading channels," *IEEE Trans. Comm.*, vol. 46, pp. 494–502, Apr. 1998.

[73] T. T. Tjhung and C. C. Chai, "Fade statistics in Nakagami-lognormal channels," *IEEE Trans. Comm.*, vol. 47, pp. 1769–1772, Dec. 1999.

[74] M. J. Barrett, "Error probability for optimal and suboptimal quadratic receivers in rapid Rayleigh fading channels," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, pp. 302–304, Feb. 1987.

[75] C. C. Tan and N. C. Beaulieu, "Infinite series representations of the bivariate Rayleigh and Nakagami-m distributions," *IEEE Trans. Comm.*, vol. 45, pp. 1159–1161, Oct. 1997.

[76] K.-W. Yip and T.-S. Ng, "Karhunen-Loeve expansion of the WSSUS channel output and its application to efficient simulation," *IEEE Trans. Comm.*, vol. 15, pp. 640–646, May 1997.

[77] F. Babich, O. E. Kelly, and G. Lombardi, "Generalized markov modeling for flat fading," *IEEE Trans. Comm.*, vol. 48, pp. 547–551, Apr. 2000.

[78] P. G. Babalis and C. N. Capsalis, "Impact of the combined slow and fast fading channel characteristics on the symbol error probability for multipath dispersionless channel characterized by a small number of dominant paths," *IEEE Trans. Comm.*, vol. 47, pp. 653–657, May 1999.

[79] G. Taricco, E. M. Biglieri, and G. Caire, "Impact of channel-state information of coded transmission over fading channels with diversity reception," *IEEE Trans. Comm.*, vol. 47, pp. 1284–1287, Sept. 1999.

[80] D.-S. Shiu, G. J. Foschini, M. J. Gans, and H. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Comm.*, vol. 48, pp. 502–513, Mar. 2000.

[81] M. Zorzi, R. R. Rao, and L. B. Milstein, "Error statistics in data transmission over fading channels," *IEEE Trans. Comm.*, vol. 46, pp. 1468–1477, Nov. 1998.

[82] M. Patzold, U. Killat, F. Laue, and Y. Li, "On the statistical properties of deterministic simulation models for mobile fading channels," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 254–269, Feb. 1998.

[83] K.-W. Yip and T.-S. Ng, "A simulation model for Nakagami-m fading channels, $m < 1$," *IEEE Trans. Comm.*, vol. 48, pp. 214–221, Feb. 2000.

[84] L. J. Mason, "Error probability evaluation for systems employing differential detection in a Rician fast fading environment and Gaussian noise," *IEEE Trans. Comm.*, vol. 35, pp. 39–46, Jan. 1987.

[85] J. Foerster and L. B. Milstein, "Analysis of hybrid, coherent FDMA/CDMA systems in Ricean multipath fading," *IEEE Trans. Comm.*, vol. 45, pp. 15–18, Jan. 1997.

[86] Q. T. Zhang, "Probability of error for equal-gain combiners over Rayleigh channels: Some closed-form solutions," *IEEE Trans. Comm.*, vol. 45, pp. 270–273, Mar. 1997.

[87] M. Shimizu, N. Aoki, K. Shirakawa, Y. Tozawa, N. Okubo, and Y. Daido, "New method of analyzing BER performance of GFSK with postdetection filtering," *IEEE Trans. Comm.*, vol. 45, pp. 429–436, Apr. 1997.

[88] T. T. Tjhung and C. C. Chai, "Error probability performance of l-branch diversity reception of MQAM in Rayleigh fading," *IEEE Trans. Comm.*, vol. 46, pp. 179–181, Feb. 1998.

[89] A. Ckalingam, M. Zorzi, L. B. Milstein, and P. Venkataram, "Performance of a wireless access protocol on correlated Rayleigh-fading channels with capture," *IEEE Trans. Comm.*, vol. 46, pp. 644–655, May 1998.

[90] M. K. Simon and M.-S. Alouini, "A unified approach to the probability of error for noncoherent and differentially coherent modulations over generalized fading channels," *IEEE Trans. Comm.*, vol. 46, pp. 1625–1638, Dec. 1998.

[91] T. K. Moon and C. Lo, "Wavelet multiscale signaling and its performance in a Ricean fast fading channel," *ITC (International Telemetering Conference)*, pp. 345–354, Oct. 1999.

[92] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels: A Unified Approach to Perform Analysis.* New York: Wiley, 2000.

[93] R. Reggiannini, "A fundamental lower bound to the performance of phase estimators over Rician-fading channels," *IEEE Trans. Comm.*, vol. 45, pp. 775–778, July 1997.

[94] E. Baccarelli, "Performance bounds and cutoff rates for data channels affected by correlated randomly time-variant multipath fading," *IEEE Trans. Comm.*, vol. 46, pp. 1258–1261, Oct. 1998.

[95] J. Mazo, "Exact matched filter bound for two-bean Rayleigh fading," *IEEE Trans. Comm.*, vol. 39, pp. 1027–1030, July 1991.

[96] F. Ling, "Matched filter-bound for time-discrete multipath Rayleigh fading channels," *IEEE Trans. Comm.*, vol. 43, pp. 710–713, Feb./Mar./Apr. 1995.

[97] K.-W. Yip and T.-S. Ng, "Matched filter bound for multipath Rician-fading channels," *IEEE Trans. Comm.*, vol. 46, pp. 441–445, Apr. 1998.

[98] C. Lo and T. K. Moon, "Matched filter bound of multiscale wavelet signaling over time-discrete multipath Rayleigh fading channels," *Proceedings of the International Telemetering Conference*, pp. 355–364, 1999.

[99] R. Visoz and E. Bejjani, "Matched filter bound for multichannel diversity over frequency-selective Rayleigh-fading mobile channels," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1832–1845, Sept. 2000.

[100] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.

[101] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2007–2019, Sept. 1999.

[102] S. V. Hanly and D. N. Tse, "The multi-access fading channel: Shannon and delay limited capacities," *Proceedings of the 33rd Allerton Conference*, 1995.

[103] E. Telatar and D. Tse, "Capacity and mutual information of wideband multipath fading channels," *to appear in IEEE Trans. Infor. Th.*

[104] D. N. Tse and S. V. Hanly, "Multiaccess fading channels - part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2796–2815, Nov. 1998.

[105] S. Shamai(Shitz) and I. Bar-David, "The empirical distribution of good codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 836–846, May 1997.

[106] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1468–1489, July 1999.

[107] P. Ligdas and N. Farvardin, "Optimizing the transmit power for slow fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 565–576, Mar. 2000.

[108] U. Hansson and T. M. Aulin, "Aspects on single symbol signaling on the frequency flat Rayleigh fading channel," *IEEE Trans. Comm.*, vol. 47, pp. 874–883, June 1999.

[109] W. M. Jang, B. R. Vojcic, and R. L. Pickholtz, "Joint transmitter-receiver optimization in synchronous multiuser communications over multipath channels," *IEEE Trans. Comm.*, vol. 46, pp. 269–278, Feb. 1998.

[110] X. Dong, N. C. Beaulieu, and P. H. Wittke, "Signaling constellations for fading channels," *IEEE Trans. Comm.*, vol. 47, pp. 703–714, May 1999.

[111] J. K. Cavers, "On the validity of the slow and moderate fading models for matched filter detection of Rayleigh fading signals," *Can. J. Elect. Comput. Eng.*, vol. 17, no. 4, pp. 183–189, 1992.

[112] P. Bello and B. Nelin, "The influence of fading spectrum on the binary error probabilities of incoherent and differentially coherent matched filter receivers," *IRE Trans. Commun. Sys.*, vol. CS-10, pp. 160–168, June 1962.

[113] D. R. Hummels and R. W. Ratcliffe, "Calculation of error probability for MSK and OQPSK systems operating in a fading multipath environment," *IEEE Trans. Veh. Technol.*, vol. VT-30, pp. 112–120, Aug. 1981.

[114] P. Bello, "Aeronautical channel characterization," *IEEE Trans. Comm.*, vol. COM-21, pp. 548–563, May 1962.

[115] F. G. Tricomi, *Integral Equations.* London and New York: Interscience Publishers, 1957.

[116] G. L. Turin, "The characteristic function of Hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, pp. 199–201, June 1960.

[117] T. K. Moon and C. Lo, "A partition of unity by sequences," *American Mathematical Monthly*, submitted for review.

[118] R. A. Wooding, "The multivariate distribution of complex normal variables," *Biometrica*, vol. 43, pp. 212–215, 1956.

[119] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C.* New York: Cambridge University Press, 2 ed., 1992.

[120] S. Haykin, *Adaptive Filter Theory.* Upper Saddle River, NJ: Prentice Hall, 3 ed., 1996.

[121] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing.* Upper Saddle River, NJ: Prentice Hall, 2000.

[122] J. W. Demmel, *Applied Numerical Linear Algebra.* Philadelphia, PA: Society for Industrial and Applied Mathematics, 1997.

[123] A. Ambrosetti and G. Prodi, *A Primer of Nonlinear Analysis.* Cambridge, New York, Oakleigh: Cambridge University Press, 1993.

[124] L. Schwartz, *Cours d'Analyse.* Paris: Hermann, 1967.

Appendixes

Appendix A
Derivative of a Functional with Respect to a Function

Let function $u \in \mathcal{L}_2$ such that $u : \mathbb{R} \to \mathbb{R}$, and let functional $F : \mathcal{L}_2 \to \mathbb{R}$ such that $F(u) \in \mathbb{R}$. Then $\frac{dF(u_0)}{du} \in \mathcal{L}_2$, the derivative of F with respect to u at $u_0 \in \mathcal{L}_2$ can be defined as a linearization of F at u_0 in Gâteaux sense [123] as follows:

$$\begin{aligned} F(u_0 + \epsilon U) =& F(u_0) + \langle \frac{dF(u_0)}{du}, \epsilon U \rangle + o(\epsilon) \\ =& F(u_0) + \epsilon \langle \frac{dF(u_0)}{du}, U \rangle + o(\epsilon), \end{aligned} \tag{A.1}$$

where $\epsilon \in \mathbb{R}$, and $U \in \mathcal{L}_2$ is any function, and

$$\langle \frac{dF(u_0)}{du}, U \rangle = \int (\frac{dF(u_0)}{du})(t) U(t)\, dt. \tag{A.2}$$

As $\epsilon \to 0$, $o(\epsilon) \to 0$, from (A.1) and (A.2), we have

$$\int (\frac{dF(u_0)}{du})(t) U(t)\, dt = \lim_{\epsilon \to 0} \frac{F(u_0 - \epsilon U) - F(u_0)}{\epsilon}. \tag{A.3}$$

Assume $F(u)$ is differentiable at u_0 and $F(u_0)$ is a local maximum of F. Then for $\epsilon > 0$ and when ϵ is small, there exists $K \in \mathbb{R}$ such that

$$\frac{F(u_0 - \epsilon U) - F(u_0)}{-\epsilon} < K \int (\frac{dF(u_0)}{du})(t) U(t)\, dt < \frac{F(u_0 - \epsilon U) - F(u_0)}{\epsilon}. \tag{A.4}$$

As (A.4) is true for all $U \in \mathcal{L}_2$, especially for $U = 0$, we have

$$\int (\frac{dF(u_0)}{du})(t) U(t)\, dt = 0. \tag{A.5}$$

Additionally, (A.5) implies

$$\frac{dF(u_0)}{du} = 0, \tag{A.6}$$

since it is true for all $U \in \mathcal{L}_2$. The same argument with $\epsilon < 0$ can show that $\frac{dF(u_0)}{du} = 0$ for $F(u_0)$ being a local minimum.

The definition of derivative of a functional with respect to a function as in (A.3) can be extended to complex functions. Let $h \in \mathcal{L}_2$ such that $h : \mathbb{R} \to \mathbb{C}$, and functional $F : \mathcal{L}_2 \to \mathbb{R}$ such that $F(h) \in \mathbb{R}$. The extension may be done in the following manner. Let $h(t) = u(t) + jv(t)$, where both $u(t)$ and $v(t)$ are real, we have [1]

$$\frac{dF(h)}{dh} = \frac{1}{2}\left(\frac{\partial F(h)}{\partial u} - j\frac{\partial F(h)}{\partial v}\right), \tag{A.7}$$

and

$$\frac{dF(h)}{dh^*} = \frac{1}{2}\left(\frac{\partial F(h)}{\partial u} + j\frac{\partial F(h)}{\partial v}\right), \tag{A.8}$$

where $\frac{\partial F(h)}{\partial u}$ and $\frac{\partial F(h)}{\partial v}$ are computed as, for $\epsilon \in \mathbb{R}$, and $U, V \in \mathcal{L}_2$ being real valued,

$$\begin{aligned}\langle \frac{\partial}{\partial u}F(h)\Big|_{h_0}, U\rangle &= \int (\frac{\partial F(h_0)}{\partial u})(t)U(t)\ dt \\ &= \lim_{\epsilon\to 0}\frac{F(h_0+\epsilon U) - F(h_0)}{\epsilon} \\ &= \lim_{\epsilon\to 0}\frac{F((u_0+\epsilon U)+jv_0) - F(h_0)}{\epsilon},\end{aligned} \tag{A.9}$$

and

$$\begin{aligned}\langle \frac{\partial}{\partial v}F(h)\Big|_{h_0}, V\rangle &= \int (\frac{\partial F(h_0)}{\partial v})(t)V(t)\ dt \\ &= \lim_{\epsilon\to 0}\frac{F(h_0+j\epsilon V) - F(h_0)}{\epsilon} \\ &= \lim_{\epsilon\to 0}\frac{F(u_0+j(v_0+\epsilon V)) - F(h_0)}{\epsilon}.\end{aligned} \tag{A.10}$$

To demonstrate the above computation with an example, suppose $H(h) = \int h(t)h^*(t)\ dt$, then

$$\begin{aligned}&H((u_0+\epsilon U)+jv_0) - H(h_0) \\ =&\int [(u_0(t)+\epsilon U(t)) + jv_0(t)][(u_0(t)+\epsilon U(t)) - jv_0(t)]\ dt - H(h_0) \\ =&\int \{(u_0(t)+\epsilon U(t))^2 + v_0^2(t)\}\ dt - \int \{u_0^2(t)+v_0^2(t)\}\ dt \\ \approx&2\epsilon\int u_0(t)U(t)\ dt,\end{aligned} \tag{A.11}$$

[1] Following Haykin [120], whom defines in this manner for the complex derivatives with discrete vectors, which in turn follows Schwartz [124].

so,

$$\frac{\partial F(h_0)}{\partial u})(t) = 2u_0(t). \tag{A.12}$$

Also,

$$\begin{aligned} & H(u_0 + j(v_0 + \epsilon V)) - H(h_0) \\ =& \int [u_0(t) + j(v_0(t) + \epsilon V(t))][u_0(t) - j(v_0(t) + \epsilon V(t))]\, dt - H(h_0) \\ =& \int \{u_0^2(t) + (v_0(t) + \epsilon V(t))^2\}\, dt - \int \{u_0^2(t) + v_0^2(t)\}\, dt \\ \approx& 2\epsilon \int v_0(t) V(t)\, dt, \end{aligned} \tag{A.13}$$

gives

$$\frac{\partial F(h_0)}{\partial v})(t) = 2v_0(t). \tag{A.14}$$

That is,

$$\left.\frac{d\,F(h)}{d\,h}\right|_{h_0}(t) = h_0^*(t) \qquad \text{and} \qquad \left.\frac{d\,F(h)}{d\,h^*}\right|_{h_0}(t) = h_0(t). \tag{A.15}$$

Appendix B
Inverse of a Symmetric Kernel

Let $k(\lambda,\tau) = \mathcal{I}[g(\lambda,\tau)]$, where $\mathcal{I}(\cdot)$ stands for the inverse of the operator $g(\lambda,\tau)$. So that we have

$$\int k(\lambda,\tau)g(\tau,\nu)d\tau = \delta(\lambda-\nu). \tag{B.1}$$

For the case $g(\lambda,\tau) = g(\lambda-\tau)$ and $k(\lambda,\tau) = k(\lambda-\tau)$, (B.1) becomes

$$\int k(t-u)g(u)du = \delta(t) \tag{B.2}$$

where $u = \lambda - \tau$. Observe that the left-hand side of (B.2) represents the convolution between $k(t)$ and $g(t)$. By performing the Fourier transform to both left hand side and right hand side of (B.2), we obtain

$$K(\omega)G(\omega) = 1, \tag{B.3}$$

or

$$K(\omega) = \frac{1}{G(\omega)}, \tag{B.4}$$

where $K(\omega)$ and $G(\omega)$ are the Fourier transforms of $k(t)$ and $g(t)$, respectively. If $G(\omega) = 0$ for some ω and is bounded below, we use $K(\omega) = \frac{1}{G(\omega)+I}$ as the approximation of the inverse operator kernel, with I being some constants such that $G(\omega) + I > 0$.

Appendix C
Programs

Matlab Codes

Algorithm 4

```
function [si,SH,EV,flag] = alg(signal_length,deltaT,Bd,beta,xi,omega,K,mu,N)
% [si,SH,EV,flag] = alg(signal_length,deltaT,Bd,beta,xi,omega,K,mu,N)
%
% This program implements Algorithm 2 in the dissertation which having
% Algorithm 1 as a special case when beta = 0
%
% Output:
%
% Si: initial signal
% SH: final S, and H pair
% EV: eigenvalue recorded through out the process
% flag: flag = 1 when there is any decrease in ev
%
%
% Input:
%
% signal_length: number of points in the signal.
% deltaT : sampling time
% Bd: Doppler freq
% beta: weight number [0,1]
%        xi = [ phi_1  xi_1]
%             [ phi_2  xi_2]
%             [ phi_3  xi_3]
%             [ phi_4  xi_4]
%                    :
%             phi_i is the delay, and xi_i is the channel gain for path i
% omega: bandwidth  in angular speed
% K: span of the random variable
% mu: synchronization delay
%     0, delta
%     1, uniform
%     2, Gaussian
% N: number of iterations
%

sxi = xi;
smxi = max(sxi(:,1));
hxi=[-1*xi(:,1) xi(:,2)];
hmxi = min(hxi(:,1));
```

```
EV = [];
oldev =0;
flag = 0;
SH =[];
r = Cal_r(beta,Bd,signal_length,deltaT);% autocorrelation function of the channel
r1 = Cal_r(0,Bd,signal_length,deltaT);
W = Cal_W(beta,omega,signal_length,deltaT); % 2beta w sinc(2pi w(tau-lambda))
W0 = zeros(size(W));
h = zeros(signal_length,1);
Z1 = floor(signal_length/4);
Z = zeros(Z1,1);
s = [Z;randn((signal_length-2*Z1),1);Z]; % leave room for SPmu and shifts
s = normal_v(s,deltaT);
si = s;
for n = 1:N,
    fprintf('\nNumber of iteration:%d\n',n);
    s = SPmu(s,K,mu);
    [h,hev] = Maxeig_p_method(r,s,sxi,smxi,W,0);
    h = normal_v(h,deltaT);
    hev = h'*Rxb(r1,s,sxi,smxi,W0,0,h);
    if hev < oldev,
      flag = 1
    end
    oldev = hev;
    h = SPmu(h,K,mu);
    [s,sev] = Maxeig_p_method(r,h,hxi,hmxi,W,0);
    s = normal_v(s,deltaT);
    sev = s'*Rxb(r1,h,hxi,hmxi,W0,0,s);
    if sev < oldev,
      flag = 1
    end
    oldev = sev;
    EV = [EV; sev  hev];
    SH = [s h];
end
EV = EV*deltaT*deltaT;
```

Algorithm 4 for Non-Fading Channel

```
function [si,SH,EV,flag] =
alg_slow(signal_length,deltaT,Bd,beta,xi,omega,K,mu,N)
% [si,SH,EV,flag] = alg_slow(signal_length,deltaT,Bd,beta,xi,omega,K,mu,N)
%
% This program implements Algorithm 2 for the case when there is no fading.
%
% Output:
%
% Si: initial signal
% SH: final S, and H pair
% EV: eigenvalue recorded through out the process
% flag: flag = 1 when there is any decrease in ev
```

```
%
%
% Input:
%
% signal_length: number of points in the signal.
% deltaT : sampling time
% Bd: Doppler freq
% beta: weight number [0,1]
%       xi = [ phi_1  xi_1]
%            [ phi_2  xi_2]
%            [ phi_3  xi_3]
%            [ phi_4  xi_4]
%                  :
%            phi_i is the delay, and xi_i is the channel gain for path i
% omega: bandwidth  in angular speed
% K: span of the random variable
% mu: synchronization delay
%     0, delta
%     1, uniform
%     2, Gaussian
% N: number of iterations
%

sxi = xi;
smxi = max(sxi(:,1));
hxi=[-1*xi(:,1) xi(:,2)];
hmxi = min(hxi(:,1));
EV = [];
oldev =0;
flag = 0;
SH =[];
r = ones(signal_length,1);
W0 = zeros(signal_length,1);
h = zeros(signal_length,1);
Z1 = floor(signal_length/4);
Z = zeros(Z1,1);
s = [Z;randn((signal_length-2*Z1),1);Z]; % leave room for SPmu and shifts
s = normal_v(s,deltaT);
si = s;
for n = 1:N,
    fprintf('\nNumber of iteration:%d\n',n);
    s = SPmu(s,K,mu);
    [h,hev] = Maxeig_p_method(r,s,sxi,smxi,W0,0);
    h = normal_v(h,deltaT);
    hev = h'*Rxb(r,s,sxi,smxi,W0,0,h);
    if hev < oldev,
      flag = 1
    end
    oldev = hev;
    h = SPmu(h,K,mu);
    [s,sev] = Maxeig_p_method(r,h,hxi,hmxi,W0,0);
    s = normal_v(s,deltaT);
```

```
    sev = s'*Rxb(r,h,hxi,hmxi,W0,0,s);
    if sev < oldev,
      flag = 1
    end
    oldev = sev;
    EV = [EV; sev  hev];
    SH = [s h];
end
EV = EV*deltaT*deltaT;
```

Algorithm 5

```
function [MaxSH,pe30] =
algpairpe30(signal_length,deltaT,Bd,beta,xi,omega,K,mu,N1,N2)
% [MaxSH,pe30] = algpairpe30(signal_length,deltaT,Bd,beta,xi,omega,K,mu,N1,N2)
%
%
% Output:
%
%
%
% Input:
%
% signal_length: number of points in the signal.
% deltaT : sampling time
% beta: weight number [0,1]
%        xi = [ phi_1  xi_1]
%             [ phi_2  xi_2]
%             [ phi_3  xi_3]
%             [ phi_4  xi_4]
%                   :
%             phi_i is the delay, and xi_i is the channel gain for path i
% omega: bandwidth  in angular speed
% K: span of the random variable
% mu: synchronization delay
%     0, delta
%     1, uniform
%     2, Gaussian
% N1: number of time using random initial vector
% N2: number of iterations for each initial vector

sxi = xi;
smxi = max(sxi(:,1));
hxi=[-1*xi(:,1) xi(:,2)];
hmxi = min(hxi(:,1));
EV = [];
oldev =0;
flag = 0;
SH =[];
r = Cal_r(beta,Bd,signal_length,deltaT);%autocorrelation function of the channel
r0 = Cal_r(0,Bd,signal_length,deltaT);
```

```
W = Cal_W(beta,omega,signal_length,deltaT); %2bw sinc(2w(tau-lambda))
W0 = zeros(size(W));
if beta == 0
IW = W;            %doesn't matter
else
  IW = -9.091*W/beta;
  IW(1) = IW(1) + 61500620; % 10*(1.626e-7)^(-1);
  IW = beta*IW;
end

for n = 1:N1,
    fprintf('\nNumber of iteration:%d\n',n);
    s = zeros(signal_length,1);
    Z = zeros(Z1,1);
    h = randn((signal_length),1); %
    h = normal_v(h,deltaT);
    h = SPmu(h,K,mu);
    [s0,e01] = Maxeig_p_method(r,h,hxi,hmxi,W,0);
    s0 = normal_v(s0,deltaT);
    [s1,e02] = Mineig_cg_method(r,h,hxi,hmxi,IW,10);
    s1 = normal_v(s1,deltaT);
    mperr = 1;

    for n2 = 1:N2
      fprintf('\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b');
      fprintf('\b\b\b\b\b\b\b\b\b\b\b\b\b\bN1:%d  N2:%d Mpe:%f',n,n2,mperr);
      s = SPmu(s0,K,mu);
      [h00,eh0max] = Maxeig_p_method(r,s,sxi,smxi,W,0);
      [h10,eh0min] = Mineig_cg_method(r,s,sxi,smxi,IW,10);

      s = SPmu(s1,K,mu);
      [h11,eh1max] = Maxeig_p_method(r,s,sxi,smxi,W,0);
      [h01,eh1min] = Mineig_cg_method(r,s,sxi,smxi,IW,10);

      h0_new = h00 + h01;
      h1_new = h10 + h11;
      h0 = normal_v(h0_new,deltaT);
      h1 = normal_v(h1_new,deltaT);

      h = SPmu(h0,K,mu);
      [s00,es0max] = Maxeig_p_method(r,h,hxi,hmxi,W,0);
      [s10,es0min] = Mineig_cg_method(r,h,hxi,hmxi,IW,10);

      h = SPmu(h1,K,mu);
      [s11,es1max] = Maxeig_p_method(r,h,hxi,hmxi,W,0);
      [s01,es1min] = Mineig_cg_method(r,h,hxi,hmxi,IW,10);

      s0_new = s00 + s01;
      s1_new = s10 + s11;
      s0 = normal_v(s0_new,deltaT);
      s1 = normal_v(s1_new,deltaT);
      sh = [s0 s1 h0 h1];
```

```
      % keep the one with smaller P(e) at 30dB
      nperr = perr(sh,r0,sxi,smxi,W0,K,mu,30,deltaT);
      if mperr > nperr
        MaxSH(:,:,n) = sh;
        pe30(n) = nperr;
        mperr = nperr;
      end
    end
end
```

P(e) for Quadratic Receiver

```
function [err] = perr(SH,r,xiphi,mxiphi,W,K,mu,SNR,deltaT)
% [err] = perrpair(SH,R,XiPhi,SNR)
% SH = [s0 s1 h0 h1]
% xiphi = [ phi_1  xi_1]
%         [ phi_2  xi_2]
%         [ phi_3  xi_3]
%         [ phi_4  xi_4]
%               :
%         phi_i is the delay, and xi_i is the channel gain for path i
% mxiphi: max delay

DN = 2;
C0 = zeros(2,2);
C1 = zeros(2,2);
M0  = zeros(2,2);
M1  = zeros(2,2);
H  = zeros(2,2);
Q  = zeros(2,2);
s0 = normal_v(SH(:,1),deltaT);
s1 = normal_v(SH(:,2),deltaT);
h0 = normal_v(SH(:,3),deltaT);
h1 = normal_v(SH(:,4),deltaT);
s0 = s0*sign(s0'*h0);
s1 = s1*sign(s1'*h1);
s = SPmu(s0,K,mu);
M0 = [h0'*Rxb(r,s,xiphi,mxiphi,W,0,h0) h0'*Rxb(r,s,xiphi,mxiphi,W,0,h1);
      h1'*Rxb(r,s,xiphi,mxiphi,W,0,h0) h1'*Rxb(r,s,xiphi,mxiphi,W,0,h1)];
M0 = M0*deltaT^2;
s = SPmu(s1,K,mu);
M1 = [h0'*Rxb(r,s,xiphi,mxiphi,W,0,h0) h0'*Rxb(r,s,xiphi,mxiphi,W,0,h1);
      h1'*Rxb(r,s,xiphi,mxiphi,W,0,h0) h1'*Rxb(r,s,xiphi,mxiphi,W,0,h1)];
M1 = M1*deltaT^2;
H = [1 h0'*h1*deltaT; h1'*h0*deltaT 1];
P0 = 0;
P1 = 0;
N2 = 10^(-SNR/10); %N/2
C0 = M0 + N2*H;
C1 = M1 + N2*H;
xi= log(det(C0)/det(C1));
```

```
Q = inv(C1)-inv(C0);
a = eig(C0*Q);
b = eig(C1*Q);
if xi <= 0,
  for I = 1:DN,
    temp = 1;
    if a(I) < 0,
      for n = 1:DN,
        if n ~= I,
          temp = temp*(1/(1-(a(n)/a(I))));
        end
      end
      temp = temp*exp(-xi/a(I));
      P0 = P0 + temp;
    end
    temp = 1;
    if b(I) < 0,
      for n = 1:DN,
        if n ~= I,
          temp = temp*(1/(1-(b(n)/b(I))));
        end
      end
      temp = temp*exp(-xi/b(I));
      P1 = P1 + temp;
    end
  end
  err = 0.5*(1+P0-P1);
else
  for I = 1:DN,
    temp = 1;
    if a(I) > 0,
      for n = 1:DN,
        if n ~= I,
          temp = temp*(1/(1-(a(n)/a(I))));
        end
      end
      temp = temp*exp(-xi/a(I));
      P0 = P0 + temp;
    end
    temp = 1;
    if b(I) > 0,
      for n = 1:DN,
        if n ~= I,
          temp = temp*(1/(1-(b(n)/b(I))));
        end
      end
      temp = temp*exp(-xi/b(I));
      P1 = P1 + temp;
    end
  end
  err = 0.5*(1-P0+P1);
end
```

Power Method

```
function [V,ev] = Maxeig_p_method(r_alpha,s,xi,mxi,W,kk)
%  [V,ev] = Maxeig_p_method(r_alpha,s,xi,mxi,W,kk)
%
%  This function returns the largest eigenvalue with the corresponding
%  eigenvector using power method
%
%  r_alpha: R_\alpha
%  xi = [ phi_1  xi_1]
%       [ phi_2  xi_2]
%       [ phi_3  xi_3]
%       [ phi_4  xi_4]
%                  :
%       phi_i is the delay, and xi_i is the channel gain for path i
%  mxi: the largest delay
%  W  : 2\beta\omega\sinc(2\pi\omega\delta T n)
%  kk : a constant that allow us to manipulate the eigenvalues of the
%       kernel
%

V = randn(size(r_alpha));
V = V/norm(V);
V_old = V;
arg = 1;
i = 0;
while (arg > 1e-6) & (i<50)
  i = i+1;
  V = Rxb(r_alpha,s,xi,mxi,W,kk,V);
  V = V/norm(V);
  arg = norm(V-V_old);
  fprintf('\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b');
  fprintf('\b\b\b\b\b\b\b\bMaxeig_p_method:%d   %f',i,arg);
  V_old = V;
end
b = Rxb(r_alpha,s,xi,mxi,W,kk,V);
ev = norm(b) - kk;
```

Conjugate Gradient Method

```
function [V,ev] = Mineig_cg_method(r_alpha,s,xi,mxi,W,kk)
% [V,ev] = Mineig_cg_method(r_alpha,s,xi,mxi,W,kk)
%
% Using Conjugate Gradient Method. This method applies to symmetric
% matrix R which solves Rx = b.
% P.311 of James W. Demmel "Applied numerical linear algebra"
%
%  This program returns the smallest eigenvalue and the corresponding
%  eigenvector of R
%
%  r_alpha: R_\alpha
```

```
%  xi = [ phi_1  xi_1]
%       [ phi_2  xi_2]
%       [ phi_3  xi_3]
%       [ phi_4  xi_4]
%                    :
%       phi_i is the delay, and xi_i is the channel gain for path i
%  mxi: the largest delay
%  W  : 2\beta\omega\sinc(2\pi\omega\delta T n)
%  kk : a constant that allow us to manipulate the eigenvalues of the
%       kernel
%

V = zeros(size(r_alpha));
Vold = V;
b = randn(size(r_alpha));
b = b/norm(b);
r = b;
p = b;
flag = 1;
th = 1e-5;
arg = 1;
i = 0;
while (i<30) & (arg > 1e-5)
  i = i + 1;
  k=0;
  while flag,
    k=k+1;
    z = Rxb(r_alpha,s,xi,mxi,W,kk,p);
    v = (r'*r)/(p'*z);
    V = V + v*p;
    rnew = r - v*z;
    temp = norm(rnew);
    fprintf('\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b');
    fprintf('\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\b\bMineig_cg_method:');
    fprintf('%d  %d  %f',i,k,temp);
    if temp < th
       flag = 0;
    end
    mu = (rnew'*rnew)/(r'*r);
    r = rnew;
    p = r + mu*p;
  end
  V = V/norm(V);
  arg = norm(V-Vold);
  Vold = V;
  th = th/1.25;
  r = V;
  p = V;
  flag = 1;
end
y = Rxb(r_alpha,s,xi,mxi,W,kk,V);
ev = norm(y)-kk;
```

Land-Mobile Model Autocorrelation Function

```
function r = Cal_r(beta,Bd,n,deltaT)
% r = Cal_r(beta,Bd,n,deltaT)
%
% This program calculates the land-mobile fading channel autocorrelation
% coeffs.
%
%
% Run the composite mapping algorithm to make a positive sequence
% Modified from Dr Moon's COMPMAP.m

n=n/2;
N = 4*n;   q = n-1;   t = (0:q)';
arg = 2*pi*Bd*deltaT
h = besselj(0,arg*t);
x = [h; zeros(N-2*q-1,1); h(q+1:-1:2)];  % Conjugate even extension

% Run the composite-mapping algorithm
converged = 0;   numiter = 0;   maxiter = 2000;
while (numiter < maxiter)
  numiter = numiter+1;
  X = real(fft(x));
  idx = X < 0;
  X(idx) = zeros(size(X(idx)));       % Enforce Property 2: positive sequence
  newx = real(ifft(X));
  newx(q+2:N-q) = zeros(N-2*q-1,1);  % Enforce Property 1: length 2q+1
  if(norm(x-newx) < 1.e-10)           % check for convergence
    break;
  end
  x = newx;
end
h=x(1:(q+1));
h=h/h(1);
a = zeros(n,1);
r=[h;a];
r=r*(1-beta);
```

Frequency Containment

```
function h = Cal_W(beta,omega,n,deltaT)
% h = Cal_W(beta,omega,n,deltaT)
%
% This program calculates the coeffs of sinc function for freq
% containment.
%
% h = 2\beta\omega\sinc(2\pi\omega\delta T n)
%
% Run the composite mapping algorithm to make a positive sequence
% Modified from Dr Moon's COMPMAP.m
```

```
N = 4*n;   q = n-1;   t = (0:q)';
arg1 = 2*beta*omega;
arg2 = 6.28319*omega*deltaT;
h = sinc(arg2*t);
x = [h; zeros(N-2*q-1,1); h(q+1:-1:2)];  % Conjugate even extension

% Run the composite-mapping algorithm
converged = 0;   numiter = 0;   maxiter = 2000;
while (numiter < maxiter)
  numiter = numiter+1;
  X = real(fft(x));
  idx = X < 0;
  X(idx) = zeros(size(X(idx)));      % Enforce Property 2: positive sequence
  newx = real(ifft(X));
  newx(q+2:N-q) = zeros(N-2*q-1,1); % Enforce Property 1: length 2q+1
  if(norm(x-newx) < 1.e-10)          % check for convergence
    break;
  end
  x = newx;
end
h=x(1:(q+1));
h = h/h(1);
h = arg1*h;
```

Synchronization Delay Distribution

```
function S = SPmu(s,K,mu)
%  function S = SPmu(s,mu)
%
%  This program is used to calculate the effective signal/filter
%  according to synchronization delay distribution
%
%  mu: 0->delta, 1->uniform, 2->Gaussian
%  K: span of the random variable
%
%  S[n] = sum_k s(n-k)pmu(k)

 if mu == 0
    S = s;
 else
    N = size(s,1);
    S = zeros(N,1);
    M = (K-1)/2;
    A = zeros(M,1);
    Sa = [A;s;A];
    if mu == 1
       pmu = ones(K,1);
    else
       k = 1:K;
         pmu = gaussian_pdf((7.38/K)*(k-M-1));
    end
```

```
  pmu = pmu/sum(pmu);
  for n=1:N
    for k = 1:K
      S(n) = S(n)+Sa(n+K-k)*pmu(k);
    end
  end
end
```

Gaussian Density Function

```
function pdf = gaussian_pdf(x)
% pdf = gaussian_pdf(x)
%
% This function returns the value of Gaussian distribution
% with zero mean and unit variance
%

[r, c] = size(x);
s = r * c;
x = reshape(x, 1, s);
pdf = zeros(1, s);
k = find(isnan(x));
if (any(k))
  pdf(k) = NaN * ones(1, length(k));
end
k = find (~isinf(x));
if (any(k))
  pdf(k) = (2 * pi)^(- 1/2) * exp(- x(k) .^ 2 / 2);
end
pdf = reshape(pdf, r, c);
```

Vector Normalization

```
function [V]=normal_v(W,deltaT)
% [V]=normal_v(W)
%
% Both V, and W are (column or row) sample vectors of a
% continuous function, where V is an normalized version
% of W such that |V| = 1 (in continuous sense)
%

D = size(W);
if D(1) == 1
  V = W./sqrt(W*W'*deltaT);
elseif D(2) ==1
  V = W./sqrt(W'*W*deltaT);
else
  V = W;
end
```

C Codes

Matrix-Vector Multiplication

```
/*************************************************************************
  v = Rxb(r_alpha,s,xi,mxi,W,k,b)

  Do the matrix_vector multiplication:
  [(1-beta)R_alpha o Psi(\tau,\lambda) + beta*w*sinc[w(\tau-\lambda)]+kI]*b
   = R*b

  We have:
        R_alpha(m,n) = r_alpha(|m-n|)
        xi = [ phi_1  xi_1]
             [ phi_2  xi_2]
             [ phi_3  xi_3]
             [ phi_4  xi_4]
                    :
             phi_i is the delay, and xi_i is the channel gain for path i

        mxi is the largest delay

        Psi(m,n) = \sum_i |xi_i|^2 s(m-\phi_i)s(n-\phi)
        W(n) = w*sinc[w(n)]


    note:    ----------> m
             |
             |
             |
             |
             n

          V(n) = \sum_m R(m,n)*b(m)
               = \sum_m[(1-beta)r_alpha(|m-n|)*
                  {\sum_i |xi_i|^2 s(m-\phi_i)s(n-\phi) }
                  + beta*W(|m-n|)+k*delta(m,n)]*b(m)


  beta and 1-beta are absorbed in r_alpha and W respectively before
  this function is called

*************************************************************************/
#include <math.h>
#include "mex.h"


/* Input Arguments */

#define ir_alph     prhs[0]
#define is          prhs[1]
```

```
#define ix          prhs[2]
#define imxi        prhs[3]
#define iW          prhs[4]
#define ik          prhs[5]
#define ib          prhs[6]


/* Output Arguments */

#define ov plhs[0]

void mexFunction( int nlhs, mxArray *plhs[],
  int nrhs, const mxArray*prhs[] )
{
  double *r_alpha,*s,*xi,*xiI,*mxi,*W,*k,*b,*v,*vn,*S,*SS;
  int i,n,m,M,N,I;
  int sshift, mphi, shift;
  double R, temp;
  mxArray *Sarray;


  r_alpha = mxGetPr(ir_alpha);
  s = mxGetPr(is);
  xi = mxGetPr(ixi);
  mxi = mxGetPr(imxi);
  W= mxGetPr(iW);
  k = mxGetPr(ik);
  b = mxGetPr(ib);

  M = mxGetM(ib);
  N = M;
  ov = mxCreateDoubleMatrix(N,1,mxREAL);
  v = mxGetPr(ov);
  for(n=0;n<N;n++) *(v+n)=0;
  I = mxGetM(ixi);
  xiI = xi+I;
  sshift = 0;

  /* Prepare for allowing shifts in s for path delays */
  if(I!=1){
    if(*mxi > 0){
      sshift = *mxi;
      shift = N+sshift;
      Sarray = mxCreateDoubleMatrix(shift,1,mxREAL);
      S = mxGetPr(Sarray);
      for(n=0;n<sshift;n++) *(S+n) = 0;
      for(m=0;n<shift;n++,m++) *(S+n) = *(s+m);
    }
    else{
      shift = N - *mxi;
      Sarray = mxCreateDoubleMatrix(shift,1,mxREAL);
      S = mxGetPr(Sarray);
```

```
    for(n=0;n<N;n++) *(S+n) = *(s+n);
    for(;n<shift;n++) *(S+n) = 0;
  }
}
else{
  Sarray = mxCreateDoubleMatrix(N,1,mxREAL);
  S = mxGetPr(Sarray);
  for(n=0;n<N;n++) *(S+n) = *(s+n);
}


/* Divided into three parts to avoid abs(m-n) */
for(n=0;n<N;n++){
  vn = v+n;
  for(m=0;m<n;m++){
    R = 0;
    for(i=0;i<I;i++){
      shift = sshift - *(xi+i);
      SS = S+shift;
      temp = *(xiI+i)**(SS+m);
      temp = temp**(SS+n);
      R = R + temp;
    }
    shift = n - m;
    R = R**(r_alpha+shift);
    R = R + *(W+shift);
    R = R**(b+m);
    *vn = *vn + R;
  }
}

for(n=0;n<N;n++){
  vn = v + n;
  R = 0;
  for(i=0;i<I;i++){
    shift = n + sshift - *(xi+i);
    SS = S+shift;
    temp = *SS**SS;;
    temp = temp**(xiI+i);
    R = R + temp;
  }
  R = R**(r_alpha);
  R = R + *(W);
  R = R +*k;
  R = R**(b+n);
  *vn = *vn + R;
}

for(n=0;n<N;n++){
  vn = v+n;
  for(m=n+1;m<M;m++){
    R = 0;
```

```
      for(i=0;i<I;i++){
        shift = sshift - *(xi+i);
        SS = S+shift;
        temp = *(xiI+i)**(SS+m);
        temp = temp**(SS+n);
        R = R + temp;
      }
      shift = m - n;
      R = R**(r_alpha+shift);
      R = R + *(W+shift);
      R = R**(b+m);
      *vn = *vn + R;
    }
  }

  mxDestroyArray(Sarray);

  return;
}
```